

| Result No. | Query | Score | Match  | Length | DB        | ID        | Description        |
|------------|-------|-------|--------|--------|-----------|-----------|--------------------|
|            |       |       |        |        |           |           |                    |
| C 1        | 930   | 100.0 | 1364   | 6      | AX376727  | AX376727  | Sequence           |
| C 2        | 930   | 100.0 | 309400 | 6      | AX127153  | AX127153  | Sequence           |
| C 3        | 930   | 100.0 | 325651 | 1      | AP005283  | AP005283  | Corynebact         |
| C 4        | 930   | 100.0 | 349115 | 1      | EX927156  | EX927156  | Corynebact         |
| C 5        | 924   | 99.4  | 924    | 6      | BD165429  | BD165429  | Novel pol          |
| C 6        | 924   | 99.4  | 924    | 6      | AX123312  | AX123312  | Sequence           |
| C 7        | 439   | 47.2  | 439    | 6      | AX376729  | AX376729  | Sequence           |
| C 8        | 399   | 42.9  | 302070 | 1      | AP005223  | AP005223  | Corynebact         |
| C 9        | 146   | 15.7  | 110000 | 1      | AP006618  | AP006618  | Continuation (16 o |
| C 10       | 146   | 15.7  | 110000 | 1      | AP006618  | AP006618  | Continuation (17 o |
| C 11       | 132.8 | 14.3  | 7646   | 1      | AF116906  | AF116906  | Rhodococc          |
| C 12       | 132.8 | 14.3  | 80609  | 1      | AF116907  | AF116907  | Rhodococc          |
| C 13       | 132.8 | 14.3  | 80610  | 1      | AP001204  | AP001204  | Rhodococc          |
| C 14       | 130.2 | 14.0  | 302325 | 1      | AE017236  | AE017236  | Mycobacte          |
| C 15       | 129.6 | 13.9  | 77534  | 1      | AF235504  | AF235504  | Streptomyc         |
| C 16       | 129.6 | 13.9  | 77536  | 6      | BD235937  | BD235937  | Polyketid          |
| C 17       | 129.6 | 13.9  | 77536  | 6      | AR271638  | AR271638  | Sequence           |
| C 18       | 129.6 | 13.9  | 77536  | 6      | AE564393  | AE564393  | Sequence           |
| C 19       | 121.2 | 13.0  | 292100 | 1      | SC0939121 | SC0939121 | Streptomyc         |

```
Db 232 GTGACCATGGGCAACGACGGGAGACTCTGGAAATCGAGACCTAGCGAGCTTCATTCA 291
Qy 61 GTGCGTCAATCAGGCACTCTACCGAAACTCGCGAAAGATTAGGATCCCGAGGCCACA 120
Db 292 GTGCGTCAATCAGGCACTCTACCGAAACTCGCGAAAGATTAGGATCCCGAGGCCACA 351
Qy 121 CTTTCCAGCAATCAGCGAGTGGAAACACAGCAGGACCCCACTTTTCGACGGGCC 180
Db 352 CTTTCCAGCAATCAGCGAGTGGAAACACAGCAGGACCCCACTTTTCGACGGGCC 411
Qy 181 GGGCCCAAACTCGTCTCAACCAAGAGGCGACGCTTCTCAACCAAGCGGCGGCATC 240
Db 412 GGGCCCAAACTCGTCTCAACCAAGAGGCGACGCTTCTCAACCAAGCGGCGGCATC 471
Qy 241 GTGCGAGAAATCAACTCGCGCGCAACTGAATCAAAACGCTCATGGACCCAGAAAAGGC 300
Db 472 GTGCGAGAAATCAACTCGCGCGCAACTGAATCAAAACGCTCATGGACCCAGAAAAGGC 531
Qy 301 ACAATCCGACTGAGCTTCATGCACTCTTGGGCACTTGGATGGTCCCGAACTTATCGA 360
Db 532 ACAATCCGACTGAGCTTCATGCACTCTTGGGCACTTGGATGGTCCCGAACTTATCGA 591
Qy 361 ACAATCCGCGCGCAACCCCAACGTAGAAATTCCAACTCCACCAAGCGCGCAATGCTC 420
Db 592 ACAATCCGCGCGCAACCCCAACGTAGAAATTCCAACTCCACCAAGCGCGCAATGCTC 651
Qy 421 GTGCTAGATCGTGTGTTTGGCTGATGAACCTGACCTCGCATTTAGTTGGGCGCCAACTGCC 480
Db 652 GTGCTAGATCGTGTGTTTGGCTGATGAACCTGACCTCGCATTTAGTTGGGCGCCAACTGCC 711
Qy 481 GAGGTTGGTACCTCTTTAGGTTGGGCGCCACTGCTTCGTCAACGACTTGCCTAGCTGT 540
Db 712 GAGGTTGGTACCTCTTTAGGTTGGGCGCCACTGCTTCGTCAACGACTTGCCTAGCTGT 771
Qy 541 CCGCGAGATCACCGGCTTGCTCTCTTTCTGCGCAAGGAGAAATGCGGTTGATTAATCGG 600
Db 772 CCGCGAGATCACCGGCTTGCTCTCTTTCTGCGCAAGGAGAAATGCGGTTGATTAATCGG 831
Qy 601 GCGGAAGAACTTGTGCGGATCGAGCAGGTTTCGGCACCGGACTCTCATGATGCA 660
Db 832 GCGGAAGAACTTGTGCGGATCGAGCAGGTTTCGGCACCGGACTCTCATGATGCA 891
Qy 661 TTAGCGGAAGAACGGTTTGTCTCCAAATGTTGTTTTCGAATCCATGGAATCAACACC 720
Db 892 TTAGCGGAAGAACGGTTTGTCTCCAAATGTTGTTTTCGAATCCATGGAATCAACACC 951
Qy 721 GTGCGAGGCTTGTGAGCGCAGGTTCTGCGGCTTGGTGTGTTCCGATGGATGATCCGTAC 780
Db 952 GTGCGAGGCTTGTGAGCGCAGGTTCTGCGGCTTGGTGTGTTCCGATGGATGATCCGTAC 1011
Qy 781 CTTCCACAGTGGGAATCGTCAACCGCCACTTAGTCCACCGCTTATAGGAATAGGT 840
Db 1012 CTTCCACAGTGGGAATCGTCAACCGCCACTTAGTCCACCGCTTATAGGAATAGGT 1071
Qy 841 TTGGTGTGCGACTCAACCGGGCGCGCACTGCGGTGGATAAATCTTCGGAAGTTTCGTG 900
Db 1072 TTGGTGTGCGACTCAACCGGGCGCGCACTGCGGTGGATAAATCTTCGGAAGTTTCGTG 1131
Qy 901 GCGGGATCGAGGTATGCATTAGAAGAGGC 930
Db 1132 GCGGGATCGAGGTATGCATTAGAAGAGGC 1161
```

```
RESULT 2
AX127153/c 309400 bp DNA linear PAT 11-MAY-2001
LOCUS
DEFINITION
Sequence 7069 from Patent EP1108790.
ACCESSION
AX127153 AX114121
VERSION
AX127153.1 GI:14041141
KEYWORDS
Corynebacterium glutamicum
SOURCE
Corynebacterium glutamicum
ORGANISM
```

```
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
1
REFERENCE
AUTHORS
Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
TITLE
Novel polynucleotides
JOURNAL
Patent: EP 1108790-A 7069 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
Location/Qualifiers
source
1..309400
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/note="Seq 1 to long (3.309.400) split in 11, seq 7069
3.000.001 3.309.400 309.400"
ORIGIN
Query Match 100.0%; Score 930; DB 6; Length 309400;
Best Local Similarity 100.0%; Pred. No. 8.9e-212;
Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGACCATGGGCAACGACGGGAGACTCTGGAAATCGAGACCTAGCGAGCTTCATTCA 60
Db 128423 GTGACCATGGGCAACGACGGGAGACTCTGGAAATCGAGACCTAGCGAGCTTCATTCA 128364
Qy 61 GTGCGTCAATCAGGCGCACCTCACCGAAACTGCGAAAGATTAGGCATCCCGAGGCCACA 120
Db 128363 GTGCGTCAATCAGGCGCACCTCACCGAAACTGCGAAAGATTAGGCATCCCGAGGCCACA 128304
Qy 121 CTTTCCAGAGAAATCAGCGAGTGGAAAAACACGACGAGCACCCCACTTTTCGACGGGCC 180
Db 128303 CTTTCCAGAGAAATCAGCGAGTGGAAAAACACGACGAGCACCCCACTTTTCGACGGGCC 128244
Qy 181 GGGCGAAACTCGTCTCAACCAAGAGGCGCACGCTTCTCAACCAAGCGGCGGCATC 240
Db 128243 GGGCGAAACTCGTCTCAACCAAGAGGCGCACGCTTCTCAACCAAGCGGCGGCATC 128184
Qy 241 GTGCGAGAAATCAAACTCGCGCAACTGAAATCAAAAGCGCTCATGGACCCAGAAAAAGGC 300
Db 128183 GTGCGAGAAATCAAACTCGCGCAACTGAAATCAAAAGCGCTCATGGACCCAGAAAAAGGC 128124
Qy 301 ACAATCCGACTGGAATTCATGCACTCTTGGGCACTTGGATGGTCCCGAACTTATCGA 360
Db 128123 ACAATCCGACTGGAATTCATGCACTCTTGGGCACTTGGATGGTCCCGAACTTATCGA 128064
Qy 361 ACATTCGCGCGCAACCCCAAGTAGAAATTCAACTCCACCAAGCGGCGAGCAATGCTC 420
Db 128063 ACATTCGCGCGCAACCCCAAGTAGAAATTCAACTCCACCAAGCGGCGAGCAATGCTC 128004
Qy 421 CTGCTAGATCGTGTGTTTGGCTGATGAAACTGACCTCGCATTTAGTGGCCCCCAACCTGCC 480
Db 128003 CTGCTAGATCGTGTGTTTGGCTGATGAAACTGACCTCGCATTTAGTGGCCCCCAACCTGCC 127944
Qy 481 GAGTTGTGTACTCTTTAGGTTGGGCGCACTGCTTCGTCAACGACTTGCCTAGCTGTT 540
Db 127943 GAGTTGTGTACTCTTTAGGTTGGGCGCACTGCTTCGTCAACGACTTGCCTAGCTGTT 127884
Qy 541 CCGCGAGATCACCGCTTGCCTCTTTTCTGGCAAGGAGAAATTCGCGTTCATTAATCTGCG 600
Db 127883 CCGCGAGATCACCGCTTGCCTCTTTTCTGGCAAGGAGAAATTCGCGTTCATTAATCTGCG 127824
Qy 601 GCGGAAGAACTTTCTGTCGCGATCGAGCAGGTTTCGGCAACCGCACTCTCATGATGCA 660
Db 127823 GCGGAAGAACTTTCTGTCGCGATCGAGCAGGTTTCGGCAACCGCACTCTCATGATGCA 127764
Qy 661 TTAGCCGAAGAGCGGTTTTGTTCCGAATCGTGTTCGAATCCATGGAATCAACACC 720
Db 127763 TTAGCCGAAGAGCGGTTTTGTTCCGAATCGTGTTCGAATCCATGGAATCAACACC 127704
Qy 721 GTGCGAGGCTTGTGTCAGCGCAGGCTCTCGGCGTGTGGTGTGGTTCGATGATGATCCGTAC 780
Db 127703 GTGCGAGGCTTGTGTCAGCGCAGGCTCTCGGCGTGTGGTGTGGTTCGATGATGATCCGTAC 127644
```



DRABWETLHNDKQIVGAOHGILSAKREGAVDKRTISTARAGLDALVKLGGAAP  
 ZSAELLENARLITVEYPAAREALASAGORNVEDNTRAVDEFKAADQELSLSKG  
 SSNEYRLQVRENIDQLGVSPRDMFAGELIDPNNAEWPFVQVRIILGGFAAEWLP  
 HGLLPRVDWNAKLAALLKFNQVVTTEYKTSFPADSLIRKVDVVFVSPFRDWNQ  
 ELGKFNIRCVTPPELSALPGQDQVITILGVKRFQOQTGDPPTTRWEXDDRRKLQDRS  
 TVRLGSDNDKVVETRETVKAGKVVQADNRNIANRAELRELEROYQASOEILKVS  
 AGLDVESADAIAELDRILLELNNTPPEATLSARHEAKKOTLARVSDLLVAQSEEV  
 ASMNKRAATELKRLESIPVAHVSEIAREVEKFLANTRRVHANVDEQIARREDL  
 DKQIDANAEULRCNQIIGILRSYIETWPNARNDLQAEPPFVGEANRLGELSLDRLL  
 AEFTAKFLGNMSTRNLQGISRLRDRARREIERIEPINASLQSFNSGADARWRN  
 IRDQSGPIVRFFQKLDATSGDLSTKQAFARYALIAEIIISKLASHDADARWRN  
 TVLDTRVRVRFLEGRDSDGATVTVDSASLGGQAKVFFCLAAALRVQLAEPGA  
 HYPVATVILDEAFDRADPAFTRQTMNVFHSFGPHVLAETPLKLIQTLGIVGSTIVV  
 SYTEPNQAGALQGNSSRSREK"

gene

complement (9534. .10172)

CDS

/gene="Cg12805"

/complement (9534. .10172)

/codon\_start=1

/transl\_table=11

/product="Hypothetical protein"

/protein\_id="BAC00199.1"

/db\_xref="GI:21325578"

/translation="MNQDLDEGDTGLTFGSRKALVOLLKGMVNALQHVWRRAITF

DQDALNAVNLFLVLDEADAGVAFTRPANGROEVLVGNKNTKAMPKVLRTETLSHF

DTLIIILRLQELTMAPPGERVTVDRREIREQVLLYRVDEERDEAKLAKRFPDAPRIV

DYSLAKTETPERFVSFALRQIFPADTVAGVRABEYKFNKAHNDGNEEEK"

complement (10169. .11656)

/gene="Cg12806"

complement (10169. .11656)

/gene="Cg12806"

/codon\_start=1

/transl\_table=11

/product="Hypothetical protein"

/protein\_id="BAC00200.1"

/db\_xref="GI:21325579"

/translation="WTVVSHALGKFRQESLESLRLSDNFPVVLA VVAOYFPQGA

AKPASELQQLSDDRVRLAREGFELPKSPDSYVSDWVKSRRFVRPGSSQGTVEVPS

ELLAVLQVQWMDPHRSISASRIEISUTQALQTLALSDSEFTAKRLAELRERDRIE

RQIEAVHAGFEFVLTQIGRVADILDLAASIADFAVRVHSLDNRKRLRQLLOP

EDSRGVLEIFRPGVDLIGDSAGRSFNSFFDVLDRERSLLIDRWIREVLGRLEAID

LDSKLRTGLYRIFROMEDASFEVNGEMTGLARSLRHYVTTEFAESRRMIQLLRDTRS

AAKAAEAGVETSLNMDTPLVRIQMDVRSIAGLKNPGERVEDLPEPVEEQLDT

EVLMEQIRASETDPELEANSVLAKQSHATIEVLEHFPATQGLASIVGLLYLAMR

DVGPIRQAVSWESDDATHRRITGWQFIRGLNSEDLAESMDK"

complement (11995. .13032)

/gene="Cg12807"

complement (11995. .13032)

/gene="Cg12807"

/codon\_start=1

/transl\_table=11

/product="Zn-dependent alcohol dehydrogenases"

/protein\_id="BAC00201.1"

/db\_xref="GI:21325580"

/translation="MTTAPQEFTHAAVVEKFGHDVTVKDIDLPKPGPHQALVKVLTSG

ICTDLHLEGWPKVPFPFPFVGHGVVEVVEVVEVVEVVEVVEVVEVVEVVEVVEV

CYVCTITRGTCNEAEYGGYQNGFQVGLVYVTRVYAAIRPDGVYDLEAAPILCAVGT

VYKALKVSTRPGQFMVYSGVGLGHIAVQVAAAMGRVIAVDIADDLKLEARKHGA

FTVNAENPDSGSAVOKYNGGAHGVLTAVHEAEPQALDMABRAGTIVENGLPPGEF

PASVNIYFKGUTIRKSLVGTQDLAEALDFFARGLIKPTVSECSLDEVNGVLDRMN

KIDGRVAIRF"

complement (13402. .13620)

/gene="Cg12808"

complement (13402. .13620)

/gene="Cg12808"

/codon\_start=1

QY 1 GTGACCATGGCAACGACGGCGGAGACCTGCGAATCGACGACCTACGACGCTTCATTCA 60  
 Db 144674 GTGACCATGGCAACGACGGCGGAGACCTGCGAATCGACGACCTACGACGCTTCATTCA 144615

QY 61 GTCGCTCAATCAGAGCCACCTCACCGAAATCGCCGAAAGATTAGGCATCCCGCAGCCACCA 120  
 Db 144614 GTCGCTCAATCAGAGCCACCTCACCGAAATCGCCGAAAGATTAGGCATCCCGCAGCCACCA 144555

QY 121 CTTTCCAGACGAATCAGCCGAGTGGAAAAACAACGAGGACACCCACATTTTTCGACCGCGCC 180  
 Db 144554 CTTTCCAGACGAATCAGCCGAGTGGAAAAACAACGAGGACACCCACATTTTTCGACCGCGCC 144495

QY 181 GGCGCGAAATCGCTCTCAACCAACGAGGACACCGCTTCTCAACGAGGCGGCGGCGGCGCATC 240  
 Db 144494 GGCGCGAAATCGCTCTCAACCAACGAGGCGGCGGCGGCGGCGGCGGCGCATC 144435

QY 241 GTCGAGAAATTCAACTCCGCGCAACTGAAATCAAAACGCTCATGGAGCCCAAGAAAAAGGC 300  
 Db 144434 GTCGAGAAATTCAACTCCGCGCAACTGAAATCAAAACGCTCATGGAGCCCAAGAAAAAGGC 144375

QY 301 ACAATCCGACTGGACTTCATGCAATCTCTTGGGCACTTGGATGGTCCCGAACTTATCCGA 360  
 Db 144374 ACAATCCGACTGGACTTCATGCAATCTCTTGGGCACTTGGATGGTCCCGAACTTATCCGA 144315

QY 361 ACATTCCGCGCGCAACACCCCAACGTAAGATTCCAACTCCACCAAGCGGCAAGCAATGCTC 420  
 Db 144314 ACATTCCGCGCGCAACACCCCAACGTAAGATTCCAACTCCACCAAGCGGCAAGCAATGCTC 144255

QY 421 CTGTAGATCTGTTTTGGCTGATGAACTGACCTCGCATTTAGTTGGCCCCCAACCTGCC 480  
 Db 144254 CTGTAGATCTGTTTTGGCTGATGAACTGACCTCGCATTTAGTTGGCCCCCAACCTGCC 144195

QY 481 GAGTTGGTACTCTTTAGGGTGGGCGCACTGCTTCTGTAACGACTTGCCTTAGCTGTT 540  
 Db 144194 GAGTTGGTACTCTTTAGGGTGGGCGCACTGCTTCTGTAACGACTTGCCTTAGCTGTT 144135

QY 541 CCCGAGATCACCGCTTCCCTCTTCTGGCCAAAGAGAAATTCGCTTGCATTACTGCG 600  
 Db 144134 CCCGAGATCACCGCTTCCCTCTTCTGGCCAAAGAGAAATTCGCTTGCATTACTGCG 144075

QY 601 GCGGAAGAACCTTTCTGGCGCATCGAGCAGGTTTCGSCACCCGACTCTCTCATGATGCA 660  
 Db 144074 GCGGAAGAACCTTTCTGGCGCATCGAGCAGGTTTCGSCACCCGACTCTCTCATGATGCA 144015

QY 661 TTAGCCGAAGAACCGCGTTTGTCCCAATGTGTTTTCGAATCCATGGAATCACCAACC 720  
 Db 144014 TTAGCCGAAGAACCGCGTTTGTCCCAATGTGTTTTCGAATCCATGGAATCACCAACC 143955

QY 721 GTCGAGGCTTGTTCAGCGAGGTCCTGGCGCTTGGTGTGGTTCGATGGATGATCCGTAC 780  
 Db 143954 GTCGAGGCTTGTTCAGCGAGGTCCTGGCGCTTGGTGTGGTTCGATGGATGATCCGTAC 143895

QY 781 CTTCCACAGTGGGAATCGTCAACGCGCCACTTAGTCCACCCGCTTATAGGAACTAGGT 840  
 Db 143894 CTTCCACAGTGGGAATCGTCAACGCGCCACTTAGTCCACCCGCTTATAGGAACTAGGT 143835

QY 841 TTGTTGGCGCACTCAACGCGGCGCGGCACTCGCGTGGATTAATTCGGAAGTTCTGTG 900  
 Db 143834 TTGTTGGCGCACTCAACGCGGCGCGGCACTCGCGTGGATTAATTCGGAAGTTCTGTG 143775

QY 901 GCGGGATCGAGGTATGCATTAGAAGAGGCG 930  
 Db 143774 GCGGGATCGAGGTATGCATTAGAAGAGGCG 143745

RESULT 4

BX927156/c

LOCUS

DEFINITION

ACCESSION

VERSION

BX927156 349115 bp DNA linear BCT 10-JUN-2004  
 Corynebacterium glutamicum ATCC 13032, IS fingerprint type 4-5,  
 complete genome; segment 9/10.  
 BX927156 BX927147  
 BX927156.1 GI:41326831

Query Match 100.0%; Score 930; DB 1; Length 325651;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-212;  
 Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|           |   |      |   |
|-----------|---|------|---|
| KEYWORDS  | complete genome.  | gene | YVRTAIGAGIPKTEVVARNVLRNALITPTTVIGLRVGYLMGGAVIIIEIFIENIQAMQOL  |
| SOURCE    | Corynebacterium glutamicum ATCC 13032                                   | CDS  | ILDGVRNDVYLQVQVTLTAIAFAIIVNIAVDLLYLVPNPRIRSI"                 |
| ORGANISM  | Corynebacterium glutamicum ATCC 13032                                   |      | 3064. .5067   |
|           | Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;            |      | /locus_tag="cg2939"   |
|           | Corynebacterineae; Corynebacteriaceae; Corynebacterium.                 |      | /locus_tag="cg2939"   |
| REFERENCE | 1 (bases 1 to 349115)   |      | /locus_tag="cg2939"   |
| AUTHORS   | Kalinowski,J., Bathe,B., Bartels,D., Bischoff,N., Bott,M.,              |      | /codon_start=1  |
|           | Burkovski,A., Dusch,N., Eggeling,L., Eikmann,B.J., Gaigalat,L.,         |      | /transl_table=11  |
|           | Goesmann,A., Hartmann,M., Huthmacher,K., Kramer,R., Linke,B.,           |      | /product="ABC-type dipeptide/oligopeptide/nickel transport    |
|           | McHardy,A.C., Meyer,F., Mockel,B., Pfeifferle,W., Puhler,A.,            |      | system, fused permease and ATPase components"                 |
|           | Rey,D.A., Ruckert,C., Rupp,O., Sahm,H., Wendisch,V.F., Wiegand,I.       |      | /protein_id="CAF20678.1"                                      |
|           | and Tauch,A.  |      | /db_xref="GI:41326834"  |
| TITLE     | The complete Corynebacterium glutamicum ATCC 13032 genome sequence      |      | /translation="MRRKLTITTTLENKPGARLGGFRALAPTISKIALVFLLLIFLLAIFA |
|           | and its impact on the production of L-aspartate-derived amino acids     |      | PIAUKYDPLASGTPVQPPSGEHWFGTDIGRDFSRVAHGASLITIGLPTGAALLV        |
|           | and vitamins  |      | AAVLGSIATAANKAISEVLMRVLDVIMSPGIALAAVFAVFGNGVPVIFATAFIYT       |
| JOURNAL   | J. Biotechnol. 104 (1-3), 5-25 (2003)                                   |      | POLARIVRANILISQFGEYVNASKVIKGASTAHILIKHVARNCAPILVATVIVADAI     |
| MEDLINE   | 22830012  |      | VFEASISFINAGVKPPSPSGWNLADGKALLSGAWPTFFFGMLILLITVCLNLISE       |
| PUBMED    | 12948626  |      | GLUTLASPKPKPVASAKALKKEESGEGSVLGHITFREANASLASLAALST            |
| REFERENCE | 2 (bases 1 to 349115)   |      | SENNNRLIFDGNPTPLLEVRDLKISFNAHGDINI VGVNFTVAPGOTMGLVSGSG       |
| AUTHORS   | Kalinowski,J.   |      | CGKSIAMIMGLPPTAKIEGILFDGKNLIDLKPDENALRGHEIAMYODALSSSL         |
| TITLE     | Direct Submission   |      | NPSMLISAKMOLTRRGKRSAAEILLELVLDPKRTLOSPHELSGQQRVLIALMAL        |
| JOURNAL   | Submitted (21-JAN-2004) Joern Kalinowski, Institut fuer                 |      | TENPRLIADDEPTTALDVTVOQVVDLLNELPEKLGPMIFVSHDLALVARLVKLTIV      |
|           | Genomforschung, Universitaet Bielefeld, Universitaetsstrasse 25,        |      | MYAGQVVGQTTREILIDPRHEYTRGLLSVLSIAGVDRLIQVPGTVPSPKEFVAGD       |
|           | 23615 Bielefeld, Germany  |      | RPAPRSEPELGDQKPVLRPITGTGTHAYATDELLAAKGEQR"                    |
|           | E-mail:Joern.kalinowski@cebitec.uni-bielefeld.de                        |      | 5064. .5882   |
|           | This sequence was accomplished by collaboration between Degussa AG      |      | /locus_tag="cg2940"   |
|           | and Bielefeld University.   |      | 5064. .5882   |
|           | Join(BX927148.1:1. .348071, BX927149.1:51. .349887,                     |      | /locus_tag="cg2940"   |
|           | EX927150.1:51. .348475, BX927152.1:51. .349799, BX927153.1:51. .349584, |      | /codon_start=1  |
|           | EX927154.1:51. .349575, BX927155.1:51. .349136, BX927156.1:51. .349115, |      | /transl_table=11  |
|           | EX927157.1:51. .140057).  |      | /product="ATPase components of ABC-type transport system,     |
| FEATURES  | Location/Qualifiers   |      | contain duplicated ATPase domains"                            |
| source    | 1. .349115  |      | /protein_id="CAF20679.1"                                      |
|           | /organism="Corynebacterium glutamicum ATCC 13032"                       |      | /db_xref="GI:41326835"  |
|           | /mol_type="genomic DNA"   |      | /translation="MTSTIDTRQPIVLVRLNTVNIHKRTGKLPDQVHANKNINPKA      |
|           | /strain="DSM 20300 = ATCC 13032"  |      | YRDEVIGIVSGSCGKSTLARMVGLQVPTSGEVLFGKPMKPRQGRKELGSSVSV         |
|           | /db_xref="taxon:196627"   |      | FQDPATSLNPRMTREQLLDPLRVHKVGDSERQWSELSI SMVLGQSALEVLPRQV       |
|           | /notes="IS fingerprint type: 4-5"                                       |      | SGGQRRVIAARALAKPDI I VADEPTSALDVSVRAQVNLNLLDLKTELGLGLVFISH    |
| gene      | 307. .1917  |      | DINTVRVSDRIAVMLAGEIIEENTTSSEIFNNAQDQYTTTLEATPSLUNKRL"         |
|           | /locus_tag="cg2937"   |      | complement(5935. .6558)                                       |
| CDS       | 307. .1917  |      | /locus_tag="cg2941"   |
|           | /locus_tag="cg2937"   |      | complement(5935. .6558)                                       |
|           | /codon_start=1  |      | /locus_tag="cg2941"   |
|           | /transl_table=11  |      | /codon_start=1  |
|           | /product="ABC-type dipeptide/oligopeptide/nickel transport              |      | /product="lyseE type translocator"                            |
|           | system, secreted component"   |      | /protein_id="CAF20680.1"                                      |
|           | /protein_id="CAF20676.1"  |      | /db_xref="GI:41326836"  |
|           | /db_xref="GI:41326832"  |      | /translation="MDAASWAFALALLVALAVPGPDLVLVLSATRGIRGTGVTAAAG     |
|           | /translation="MSTTITRRNFRATGILGVAAGIGATLAACAPDNTGTSSTSTA                |      | IMTGMLHASIAIAGATALLSAPGVLSAIOQLGAGVLLMGTNMFRASTONTGESSTA      |
|           | AGTGANEGTITAAISVELGTNGYDPMTTTSALTVAANWHTLEGLEIDPATGEVYA                 |      | ASQSSAGYFRGFI TNATNPKALLFFAAILPOFIGNGEDMKMETLALCATIVILGSGAWW  |
|           | ALASALPSADATSLDILKRDGATFHNGDAVTDVVPSERVLDPANNSLYASFIPFI                 |      | IGTIALVRGIGLOKLP SADRIITLVGGIALFLIAGLLVNTAYGLIT"              |
|           | KSVTKDDTVIDLYATGISERLAVKVIKPSVVEADASGFDPANPGISGPIKMTD                   |      | 6630. .7130   |
|           | NGAKVKVFERNDYNGPRAPAAKNWQIIPDASTRNSLQSGTMAIDSVPEYLSIP                   |      | /locus_tag="cg2942"   |
|           | QLEATSVESQVGLLFAMFNCSEGNPNFNQAFLYALDMDKIVKTNQMSDQATP                    |      | 6630. .7130   |
|           | ATSFQVKEHPNTNQASTVSLDADAKALFAETGLTSLNLTCTDHDWKNVNCPLIQES                |      | /locus_tag="cg2942"   |
|           | LAALGNVGSFTRKSADVNTTIDGPEAVDVVIA PGDPSVFGNDPDLMLRWYAGDVW                |      | /codon_start=1  |
|           | TDSPRMHKGSESVDQVQNLLEGIRATKAEQQDIWNRFTDISNVPIYPLFHRKVP                  |      | /transl_table=11  |
|           | TAWNSNALVDFRPSLUTGNLFSGVATTE"   |      | /product="Bacterial regulatory proteins, AsnC family"         |
| gene      | 2098. .3063   |      | /protein_id="CAF20681.1"                                      |
|           | /locus_tag="cg2938"   |      | /db_xref="GI:41326837"  |
| CDS       | 2098. .3063   |      | /translation="NNBFQMKODQSPKLDSDVMKLELLOVDATMPNKDIAAKVGTAP     |
|           | /locus_tag="cg2938"   |      | STCLERIRMRBNGTIVATRAHVRPSLLGRGEQAPLGIGIRPHARDTNDFFQVKVIAL     |
|           | /codon_start=1  |      | PETLALYNVSGSEDTLVHVAVANSTELQSLIIDKLALUPQVAHCRQTQLIFGEPVAPL    |
|           | /transl_table=11  |      | RQSDGQ"   |
|           | /product="ABC-type dipeptide/oligopeptide/nickel transport              |      | complement(7178. .7540)                                       |
|           | system, permease component"   |      | /locus_tag="cg2943"   |
|           | /protein_id="CAF20677.1"  |      | complement(7178. .7540)                                       |
|           | /db_xref="GI:41326833"  |      | /locus_tag="cg2943"   |
|           | /translation="MSNLRLLVGRRLIALPIMIGVTLVLFVIMSPSPADPARLALGE               |      | /codon_start=1  |
|           | SASPEALEYRANGLNDPMVRYFDILGMLKGDGLTSGSGVAVTDIVARAPFITLQ                  |      | /transl_table=11  |
|           | LPFWGLITIAVVALIGVIAALYRDRWPDQLIRVVSIALATPFSFWALLILQWLGTTI               |      | /product="putative membrane protein"                          |
|           | PGAWGFFPALVTRWVPFSEDPATYFNIALPAIALAVPVAGSLARVVRTSMVEELDKD               |      | /protein_id="CAF20682.1"                                      |

```
/db_xref="GI:41326838"
/translation="MPKLUSSPKVNNLVAIIVLGGMATEQIIIGFRDNNSTALVMWISSA
IAIGNVWGLSNYKPCDFPIRLYAFAPVNLMLLFDQAQSWTFWGGLAGAVIFTLI
LSGFALYDFHKSPPRK"
gene complement(7624..8106)
/locus_tag="cg2944"
/locus_tag="cg2944"
CDS complement(7624..8106)
/locus_tag="ispf"
/locus_tag="cg2944"
/EC_number="2.7.7.-"
/codon_start=1
/transl_table=11
/product="2-C-METHYL-D-ERYTHRITOL 4-PHOSPHATE
CYTIDYLTRANSFERASE"
protein_id="CAF20683.1"
/db_xref="GI:41326838"
/translation="MTNPIIPRVGIATDAHQIEAKPCWIACLIFEGVDGCEHSDGD
VVAHAIVALLSASGLDGFVGVGREYDVGSTOLLKEVRELLSAHGVVIGNVAA
QLVGQTPKFGPRREAQVISEIIGAPCSLSATITDHHGFTGRSEGRASVATAVWKA"
gene complement(8099..8869)
/locus_tag="cg2945"
CDS complement(8099..8869)
/locus_tag="ispf"
/locus_tag="cg2945"
/EC_number="2.7.7.-"
/codon_start=1
/transl_table=11
/product="2-C-METHYL-D-ERYTHRITOL 4-PHOSPHATE
CYTIDYLTRANSFERASE"
protein_id="CAF20684.1"
/db_xref="GI:41326840"
/translation="MSSTRIPVIALAAAGTRIGGPIPKAFVTLRERTLLERSLOA
MLTSSVDEIILVSPDMETAYARDLLRGLNDPGEVRLVHGGGRADSVMAGLQ
AISLDATPDALVLHDSARLATPPGMIAVRVVRKHEGATAVIPVLVPSVDIKRVSPD
GGVVDTNRAELRAVQTPQGFLLSELVAANEKFPADPNPGFIPITDDASLMEWYGADV
VCVQGDPAFVKTTPIDNMLAQRIIDEAEPIFEVPGD"
gene complement(8876..9472)
/locus_tag="cg2946"
CDS complement(8876..9472)
/locus_tag="cg2946"
/codon_start=1
/transl_table=11
/product="Card-like transcriptional regulator"
protein_id="CAF20685.1"
/db_xref="GI:41326841"
/translation="MEFKVGDVVYPVPHHGAALISALEQREMNMGETVDYLVQLINHSDL
100.0%; Score 930; DB 1; Length 349115;
Best Local Similarity 100.0%; Pred. No. 8.9e-212;
Matches 910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query Match
1 GTGACATGGGCAAGCAGCGCGAGACCTGCGAATCGACGACTACGAGCTTCATTCA 60
308144 GTGACATGGGCAAGCAGCGCGAGACCTGCGAATCGACGACTACGAGCTTCATTCA 308085
61 GTGCTCAATCAGGCACTCAACCGAACTGCGGAAAGATTAGGATCCCGAGGCCACA 120
308084 GTGCTCAATCAGGCACTCAACCGAACTGCGGAAAGATTAGGATCCCGAGGCCACA 308025
121 CTTTCCAGAGCAATCAGCGAGTGGAAACAAACGAGGACCCCTTTTCAGCCGCGC 180
308024 CTTTCCAGAGCAATCAGCGAGTGGAAACAAACGAGGACCCCTTTTCAGCCGCGC 307965
181 GGCAGCAACTCGTCTCAACCAACGAGGCCACCGCTTCTCAACCAACGAGGCCACATC 240
307964 GGCAGCAACTCGTCTCAACCAACGAGGCCACCGCTTCTCAACCAACGAGGCCACATC 307905
241 GTGCGAGAAATTCATCCCGCGCACTGAAATCAACGCTCATGACCCAGAAAAGGC 300
307904 GTGCGAGAAATTCATCCCGCGCACTGAAATCAACGCTCATGACCCAGAAAAGGC 307845
```

```
QY 301 ACAATCCGACTGGACTTCATGCATTCTTGGGCACTTTGGATGTCCTCCGAACTTATCCGA 360
Db 307844 ACAATCCGACTGGACTTCATGCATTCTTGGGCACTTTGGATGTCCTCCGAACTTATCCGA 307785
QY 361 ACATTCCGCGCGCAACACCCCAACAGTAGAATTCACAACTCCCAAGCGGAGCAATGCTC 420
Db 307784 ACATTCCGCGCGCAACACCCCAACAGTAGAATTCACAACTCCCAAGCGGAGCAATGCTC 307725
QY 421 CTGCTAGATCGTGTGTTGGCTGATGAACCTCGACTAGTTCGCTTGGCCCCAAACCTGCC 480
Db 307724 CTGCTAGATCGTGTGTTGGCTGATGAACCTCGACTAGTTCGCTTGGCCCCAAACCTGCC 307665
QY 481 GAGGTGTGATCTCTTTAGGGTGGCGCCACTGCTGTTCAACGACTTGGCCCTAGCTGTT 540
Db 307664 GAGGTGTGATCTCTTTAGGGTGGCGCCACTGCTGTTCAACGACTTGGCCCTAGCTGTT 307605
QY 541 CCCGAGATCACCGGCTTGCCTCTTTCTGGCCAAAGAGAAATTCGCGTTGAATCTACGCG 600
Db 307604 CCCGAGATCACCGGCTTGCCTCTTTCTGGCCAAAGAGAAATTCGCGTTGAATCTACGCG 307545
QY 601 CGGAAGAACCTTTCGTGGCGATCGGAGCAGGTTTCGGCACCCGACTCTCATGGATGCA 660
Db 307544 CGGAAGAACCTTTCGTGGCGATCGGAGCAGGTTTCGGCACCCGACTCTCATGGATGCA 307485
QY 661 TTAGCCGAAGAAGCGGTTTTGTTCCCAATGCTGTTTTCGAATCCATGGAACCTCACCAACC 720
Db 307484 TTAGCCGAAGAAGCGGTTTTGTTCCCAATGCTGTTTTCGAATCCATGGAACCTCACCAACC 307425
QY 721 GTCGAGGGCTTGTGACGCGAGGTCCTCGCGTTTGGTGTGTTTCCGATGATGATCCGTCAC 780
Db 307424 GTCGAGGGCTTGTGACGCGAGGTCCTCGCGTTTGGTGTGTTTCCGATGATGATCCGTCAC 307365
QY 781 CTTCCACAGTGGGAATCGTCAAGCCCACTTAGTCCACCGCTTATAGGGAACCTAGGT 840
Db 307364 CTTCCACAGTGGGAATCGTCAAGCCCACTTAGTCCACCGCTTATAGGGAACCTAGGT 307305
QY 841 TTGTTGTGGGACTCAACCGCGGCGCGGCACTGCGGTGGGATAACTCCGGAAGTTCGTG 900
Db 307304 TTGTTGTGGGACTCAACCGCGGCGCGGCACTGCGGTGGGATAACTCCGGAAGTTCGTG 307245
QY 901 CGCGGATCGAGGTATGCAATTAGAGAGGCG 930
Db 307244 CGCGGATCGAGGTATGCAATTAGAGAGGCG 307215
RESULT 5
BD165429
LOCUS BD165429 Novel polynucleotide. 924 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel polynucleotide.
ACCESSION BD165429
VERSION BD165429.1 GI:27871241
KEYWORDS JP 2002191370-A/3228.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 924)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.
TITLE Novel polynucleotide
JOURNAL Patent: JP 2002191370-A 3228 09-JUL-2002;
KYOWA HAKKO KOGYO CO LTD
COMMENT OS Corynebacterium glutamicum
PN JP 2002191370-A/3228
PD 09-JUL-2002
PF 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI,
PI KEIKO OCHIAI,
PI HARUHIKO YOKOI,NAOKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO
PI OZAKI
PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC
PC C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/
```

|                           |  |
|---------------------------|--|
| PC                        | 04, C12P13/08,   |
| PC                        | C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, P    |
|                           | G01N33/566,  |
| PC                        | G01N33/569, G01N33/68, G01N37/00/C12P21/08, (C12N1/21, C12R1.15),    |
| PC                        | (C12N1/21, C12R1.13), (C12N1/21, C12R1.01), (C12P13/08, C12R1.15),   |
| PC                        | C12N15/00,   |
| PC                        | C12N5/00, C12N15/00  |
| CC                        | Novel polynucleotide   |
| FH                        | Key  |
| FT                        | source   |
| FT                        | 1..924   |
| FT                        | Location/Qualifiers  |
| FT                        | Location/Qualifiers  |
| FT                        | 1..924   |
| FT                        | /organism="Corynebacterium glutamicum".                              |
| FT                        | /organism="unidentified"   |
| FT                        | /mol_type="genomic DNA"  |
| FT                        | /db_xref="taxon:32644"   |
| ORIGIN                    |  |
| Query Match               | 99.4%; Score 924; DB 6; Length 924;                                  |
| Best Local Similarity     | 100.0%; Pred. No. 2.8e-210;  |
| Matches 924; Conservative | 0; Mismatches 0; Indels 0; Gaps 0;                                   |
| Qy                        | 7 ATGGGCAACGACGGCGGAGACCTCGAATCGACGACCTACGACGTTTCATTTCAGTCGCT 66     |
| Db                        | 1 ATGGGCAACGACGGCGGAGACCTCGAATCGACGACCTACGACGTTTCATTTCAGTCGCT 60     |
| Qy                        | 67 CAATCAGGCCACCTCACCGAAATGCGGAAAGATTAGGCATCCCGCAGCCACACATTTCC 126   |
| Db                        | 61 CAATCAGGCCACCTCACCGAAATGCGGAAAGATTAGGCATCCCGCAGCCACACATTTCC 120   |
| Qy                        | 127 AGCGAATCAGCCGAGTGGAATAACACGACGAGCACCCACATTTTCGACCGCGCGCGC 186    |
| Db                        | 121 AGCGAATCAGCCGAGTGGAATAACACGACGAGCACCCACATTTTCGACCGCGCGCGC 180    |
| Qy                        | 187 AAATCTGCTCTCAACCAACGAGGCGCGCTTCCTCAACGACGCGGCGCATGTCGCA 246      |
| Db                        | 181 AAATCTGCTCTCAACCAACGAGGCGCGCTTCCTCAACGACGCGGCGCATGTCGCA 240      |
| Qy                        | 247 GAATTCAAATCCGCGGCAATCGAATCAAAACGCTCATGGACCCAGAAAGGACCAATC 306    |
| Db                        | 241 GAATTCAAATCCGCGGCAATCGAATCAAAACGCTCATGGACCCAGAAAGGACCAATC 300    |
| Qy                        | 307 CGACTGCACTTCATGCATTCCTTGGGCACTTGGATGTTCCCGAATTCATCCGAACATTC 366  |
| Db                        | 301 CGACTGCACTTCATGCATTCCTTGGGCACTTGGATGTTCCCGAATTCATCCGAACATTC 360  |
| Qy                        | 367 CGCGCCGAACACCCCAAGTGAATTCAACTCAACGACGCGGCGAGCAATGCTCTGTA 426     |
| Db                        | 361 CGCGCCGAACACCCCAAGTGAATTCAACTCAACGACGCGGCGAGCAATGCTCTGTA 420     |
| Qy                        | 427 GATCGTGTTTTGGCTGATGAACCTGACCTCGCATTTAGTTGGCCCAACCTCGCGAGTT 486   |
| Db                        | 421 GATCGTGTTTTGGCTGATGAACCTGACCTCGCATTTAGTTGGCCCAACCTCGCGAGTT 480   |
| Qy                        | 487 GGTAACCTCTTTAGGTGGCGCCACTGCTTCGTCAACGACTTGCCTAGCTGTTCCCGCA 546   |
| Db                        | 481 GGTAACCTCTTTAGGTGGCGCCACTGCTTCGTCAACGACTTGCCTAGCTGTTCCCGCA 540   |
| Qy                        | 547 GATCACCGGCTTGCTCTCTTTTCGGCCAAAGAGAAATTCGCGTGTGATTACTCGGCGGAA 606 |
| Db                        | 541 GATCACCGGCTTGCTCTCTTTTCGGCCAAAGAGAAATTCGCGTGTGATTACTCGGCGGAA 600 |
| Qy                        | 607 GAACCTTTCTGGGCGATCGGAGCAGTTTCGGCACCCGACCTCCATCGATGCAATTAGCC 666  |
| Db                        | 601 GAACCTTTCTGGGCGATCGGAGCAGTTTCGGCACCCGACCTCCATCGATGCAATTAGCC 660  |
| Qy                        | 667 GAAGAAGCCGTTTTTGTTCCTCAATGTGTTTTTCAATCCATGGAACCTCACCACCGTCGA 726 |
| Db                        | 661 GAAGAAGCCGTTTTTGTTCCTCAATGTGTTTTTCAATCCATGGAACCTCACCACCGTCGA 720 |
| Qy                        | 727 GGGCTTGTACGCGAGTCTCGGCGTTGGTTCGATGGATGATCGGTACCTTCCC 786         |
| Db                        | 721 GGGCTTGTACGCGAGTCTCGGCGTTGGTTCGATGGATGATCGGTACCTTCCC 780         |



QY 487 GGTACCTCTTTAGGGTGGCGGCACCTGCTTCGTCAACGACTTGCCTAGCTGTTCCCGCA 546  
Db 481 GGTACCTCTTTAGGGTGGCGGCACCTGCTTCGTCAACGACTTGCCTAGCTGTTCCCGCA 540  
QY 547 GATCACCGGCTTGCCTCTCTTTCTGGCCAAAGAGAAATTCGCGTTGATTAATGACGCGGAA 606  
Db 541 GATCACCGGCTTGCCTCTCTTTCTGGCCAAAGAGAAATTCGCGTTGATTAATGACGCGGAA 600  
QY 607 GAACCTTTCTGGCGATCGAGCAGGTTTCGGCACCACGACTCCTCATGATGCAATAGCC 666  
Db 601 GAACCTTTCTGGCGATCGAGCAGGTTTCGGCACCACGACTCCTCATGATGCAATAGCC 660  
QY 667 GAAGAAGCCGGTTTGTTCCTCAATGCTGTTTTTTCGAATCCATGGAATCACCACCGTCGCA 726  
Db 661 GAAGAAGCCGGTTTGTTCCTCAATGCTGTTTTTTCGAATCCATGGAATCACCACCGTCGCA 720  
QY 727 GGGCTTGTGAGCGCAGGTTCTGGCGTTGGTGTGGTTCCGATGGATGATCCGTACCTCCC 786  
Db 721 GGGCTTGTGAGCGCAGGTTCTGGCGTTGGTGTGGTTCCGATGGATGATCCGTACCTCCC 780  
QY 787 ACAGTGGGAATCGTCAAGCCGACCTTAGTCCACCGCTTATAGGAACTAGGTTTGGTG 846  
Db 781 ACAGTGGGAATCGTCAAGCCGACCTTAGTCCACCGCTTATAGGAACTAGGTTTGGTG 840  
QY 847 TGGCGACTCAACGCGGGCGCGACCTCGCGTGGGATAACTTCCGGAAGTTTCGTGGCGGGA 906  
Db 841 TGGCGACTCAACGCGGGCGCGACCTCGCGTGGGATAACTTCCGGAAGTTTCGTGGCGGGA 900  
QY 907 TCGAGGTATGCAATTAGAGAGGCG 930  
Db 901 TCGAGGTATGCAATTAGAGAGGCG 924

RESULT 7  
AX376729  
LOCUS AX376729 439 bp DNA linear PAT 18-MAR-2002  
DEFINITION Sequence 3 from Patent WO0212504.  
ACCESSION AX376729  
VERSION AX376729.1 GI:19573114  
KEYWORDS  
SOURCE Corynebacterium glutamicum  
ORGANISM Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
REFERENCE  
1 Moeckel, B., Farwick, M., Hermann, T., Kreutzer, C. and Pfeifferle, W.  
Nucleotide sequences which code for the lysr2 gene  
Patent: WO 0212504-A 3 14-FEB-2002;  
Degussa AG (DE)  
FEATURES  
Location/Qualifiers  
source  
1. 439  
/organism="Corynebacterium glutamicum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1718"  
/note="lys2int"

ORIGIN  
Query Match 47.2%; Score 439; DB 6; Length 439;  
Best Local Similarity 100.0%; Pred. No. 3.1e-94;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 CCATCGTCGCAATTCACCTCGCGCACTGAATCAACGCTCATGACCCAGAAA 295  
Db 1 CCATCGTCGCAATTCACCTCGCGCACTGAATCAACGCTCATGACCCAGAAA 60  
QY 296 AAGGCACATCCGACTGGAATTCATGCAATTCCTGGGCACTTGGATGTCCTCCGAACTTA 355  
Db 61 AAGGCACATCCGACTGGAATTCATGCAATTCCTGGGCACTTGGATGTCCTCCGAACTTA 120  
QY 356 TCCGACATTCGCGCGGACCCAGTGAATTCCACTCCCAAGCGGCGGCA 415  
Db 121 TCCGACATTCGCGCGGACCCAGTGAATTCCACTCCCAAGCGGCGGCA 180

QY 416 TGCTCTGTAGATCGGTGTTTGGCTGATGAACTGACCTCGCATTTAGTTGGCCCCAAAC 475  
Db 181 TGCTCTGTAGATCGGTGTTTGGCTGATGAACTGACCTCGCATTTAGTTGGCCCCAAAC 240  
QY 476 CTGCCGAGGTGGTACCTCTTTAGGGTGGCGCCACTGCTTCGTCACGACTTGCCTTAG 535  
Db 241 CTGCCGAGGTGGTACCTCTTTAGGGTGGCGCCACTGCTTCGTCACGACTTGCCTTAG 300  
QY 536 CTGTTCCCGCAGATCAGCGGCTTGCCTCTTTCTGGCCAGGAGAAATTCGCGTTGATTA 595  
Db 301 CTGTTCCCGCAGATCAGCGGCTTGCCTCTTTCTGGCCAGGAGAAATTCGCGTTGATTA 360  
QY 596 CTGCCGCGGAAGAACCTTTCTGTCGATGCGAGCAGGTTTCGGCACCCGACTCCTCATGG 655  
Db 361 CTGCCGCGGAAGAACCTTTCTGTCGATGCGAGCAGGTTTCGGCACCCGACTCCTCATGG 420  
QY 656 ATGCATTAGCGGAAGAAGC 674  
Db 421 ATGCATTAGCGGAAGAAGC 439

RESULT 8  
AP005223/c  
LOCUS AP005223 302070 bp DNA linear BCT 24-JUL-2003  
DEFINITION Corynebacterium efficiens YS-314 DNA, complete genome, section 10/11.  
ACCESSION AP005223 BA000035  
VERSION AP005223.1 GI:23494433  
KEYWORDS  
SOURCE Corynebacterium efficiens YS-314  
ORGANISM Corynebacterium efficiens YS-314  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
REFERENCE  
1 Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and Gojobori, T.  
Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of Corynebacterium efficiens  
Genome Res. 13 (7), 1572-1579 (2003)  
2 (bases 1 to 302070)  
Kawarabayasi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and Director-General of Biotechnology Center.  
Submitted (17-MAY-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan (E-mail: bio@nitech.go.jp, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)  
Kawarabayasi, Y. is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, 305-8566 Japan  
Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan  
Itoh, T. is at the Japan Biological Information Research Center, Koto-ku, Tokyo, 135-0064 Japan  
Yamagishi, A. is at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392 Japan  
Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co., Inc., Kawasaki, Kanagawa, 210-8681 Japan  
The other authors are at the National Institute of Technology and Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.  
Location/Qualifiers  
source  
1. 302070  
/organism="Corynebacterium efficiens YS-314"  
/mol\_type="genomic DNA"  
/strain="YS-314"  
/db\_xref="taxon:196164"  
complement(52..1248)  
/note="CE2590, similar to X89084-2|CAA61456.1| percent

CDS



```

identity: 85 in 397 aa"
/codon_start=1
/transl_table=11
/product="acetate kinase"
/protein_id="BAC19400.1"
/db_xref="GI:23494434"
/translation="WALALVLSNGSSSIKQLVNPANHAHTDDPFVSGLVEQIGERKGR
ILKTEGKVKVEAPIADHSAGLSMFDLMTHEGCGPSQVDIIAVGHRVHGGILFSA
PLIIDEIVEMIRDLIPLAPLNPANIDGIEVARILLPDVPHVAVEDFGFHSLLPAA
ALYANKOVAEYGRYVGHGTSHSYSSRVVDLMDKPABEINTITHLGNGASMAA
VKGVAVDTSMTGTPLAGVMCTRTGDIDPGVVFHLARNANSIDEIDNLNKKSGVK
GLSVDNDFRELQMLDEGDQDAMSAYNIYIHLQRLYGSYMWALGRVCLVFTAGVE
NAHFVREDALAGLEWGIKVPDERNKLPNDGPRLLISTDDSTVKVFVIPTNEIATARY
SAKPAE"
complement(1252..2766)
/notes="CE2591, similar to AE006946-9|AAK44645.1| percent
identity: 45 in 495 aa"
/codon_start=1
/transl_table=11
/product="phosphate acetyltransferase"
/protein_id="BAC19401.1"
/db_xref="GI:23494435"
/translation="WRYRASNDAKLCESQHLPRVPSGTHLVERITPISGLREERRC
MSDTPSALITTVNRSFDGFDLEALARDLEVNLPVRGDDPMSKVLAAADLLADGPAL
IYAGNAHFDAQAALGVPMMLVDRKGSHTVLTARTQAKNVGAVVAAFTABEEAKP
DKLKAARNHGLEPMSAALFENMLKRAKEQAHIVLPEGEDDRILMAHQHLEKD
ICELTILGDEPQIHSGRATGLHLEKADVINTLPDKLDEPAEQFAELRKSGITLEQ
ARETWKDISYFTGMVNHGDAGMVSQAANTTAHTIKPSFQIITKTPVASVVSIFLM
VRLGRMAFGCAVNPNTAEQGLGIAVVSATASQFGIDPRVAILSVSTGNSGTGPD
VDRADIALAEARIDIPDLKQGLQFDDAADPGVAKKMPDSVDAGHANVSTFGPDLA
GNIGYKTAQRTQHALAVGPILOGLNKPVNDLSRGATVPDINTVTAITAIQAGGK"
3072..4511
/notes="CE2592, similar to AE001908-11|AAF10073.1| percent
identity: 42 in 465 aa"
/codon_start=1
/transl_table=11
/product="putative ferredoxin/adrenodoxin reductase"
/protein_id="BAC19402.1"
/db_xref="GI:23494436"
/translation="WRFGCPPLSGRDRQPMFERKLSMSRLRVAVVCGPAGIYAS
DLUMKSDHDVDLDERPAPFGLTRYGVAPDHPRIKGLVKSLLHNVLKEQRLLGNI
EIGKDIIVDELREFDAIIFSTGQDQLRVKGLDGSGWAGFVFGYDGNPDFAR
SWLSAEKAVTGVGNVALDVARIILAKTGDLEKVTIEIPDNVYESIKNKAREVHVFGR
RQPAQKATFPLELDELDSPLNEVVNPEDIDYDAASEOARDSDSQILVCOTLQRYA
IRDPGAKHLPIHFESPEVILGEDGVKVGKTRTRLDNGNGVVGTFGEFTTWQVA
VYRAGVSGRVIDPDEDERAIIENDGSRVIDPSTGSPVTSLYATGIRKGPISLIG
NKSDAKETTEMLLADYRAGSDPAPANPETDAIIELLDQRKIPTYITWDGHLDDAER
EAGEREGREKKIYEWNDMVRFAPEYEI"
4719..5261
/notes="CE2593"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC19403.1"
/db_xref="GI:23494437"
/translation="WAMTRYFSVSNLAEISSLEVHKLYKLRVIDIFVHOKTPYASIDD
TPALPTTHILWREHEDHPTLKGCARLVPPTVSEVAAATGATLELDDAEVSLGRV
AVQSRSRGTLSAETMENARLAYEQFGPKQVLLTRAQLQEPYEGFGFTTCGTQYDE
AGVAFVMVWLKASLVRVYAA"
complement(5283..6560)
/notes="CE2594, similar to AX064627-1|CAC25553.1| percent
identity: 86 in 401 aa"
/codon_start=1
/transl_table=11
/product="putative phosphoribosylglycinamide
formyltransferase"
/protein_id="BAC19404.1"
/db_xref="GI:23494438"
/translation="MSGVPDITYSGSHMERGAQTESQLPIPEKIGTPLSPNATKWNIL
GSGELGVETIAPQLRGVEHVAVDYDNPAPQVAFHSYVIDMTDAAAREGLFTTYPK
DFPIIEIETALATDELVRIEQEGLATVVPATARTQLTMTNREGIRRLASEELGTSYGE
FCSTFEETFAAERLGPVNVKPMVSSGSGQSVTSAPDLESAMEYMSGARVGNQR
VIVEQFVEPDYEITLLTVRGIDPATGKPATWFCPTIGHRQDQDGVESWQPMWTAPA
LENARSVAARITNALGGRGVFGVELFVGDDVYFSEVSPRPHDTGLVTLATQRFSEFE
LHKAVALGLPVDVTLTSPGASAVIYGVDSPCVSYAGLAELAAVAETDVRFGKPEAF
TKRRMGVAVTAEDTATARDRATLAAAATVHG"
5584..7768
/notes="CE2595, similar to AX064083-1|CAC25282.1| percent
identity: 38 in 391 aa"
/codon_start=1
/transl_table=11
/product="putative peptidase"
/protein_id="BAC19405.1"
/db_xref="GI:23494439"
/translation="WNALHNLDDTRADREELYKWFHRNPESLSAEFNTIERITELK
NMDIKLIDVTATGAVALIENGDPVVALRGDIDALPVTQETGLDYASEAHGVMHACG
HDFHITLLISADVDFASNRDKMSGTLVAVFPAEBETASGARMVADIASVTPKPDVY
LQGHVMAOAGVISMPETAVMTSMESIRVRVFGTGTGHSMPHLGVDPIVLAAHITRL
QTVVARETSPPDFGVVTVGAINAGTKANIIPESAEILVNTRAMSVEIOSKIRTATERI
VSSECAAGCAPEPTFEYDRAPLTFNDADARSVVAEAFNQAFGEYQVTFSTASED
FVSDALGVPVTFVFWGVGEDPENAGNHSFPFFAPVIOPTLDRSLEAFITTAASAWLV
RS"
7827..8417
/notes="CE2596"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC19406.1"
/db_xref="GI:23494440"
/translation="WKTTPRPLPLPKVPKYSGRMAVTGAPAGTFAEPREWEKARGTLH
TQIKTAMSGSYGKTIAPTLQEMEKLLPRYSVSHEDAVVSELCTLVIAEDLQCGIQ
BFCVNAPELDCMKDARILEYHAATSHLSSAAALICARFDDNSPPPAIDTAMAFDAVF
TKLGYPAPHEDTRSPRYTMSPRDINDALSHFEYRYP"
complement(8727..10016)
/genes="pura"
/complement(8727..10016)
/genes="pura"
/EC_numbers="6.3.4.4"
/notes="CE2597, similar to AL583918-8|CAC29788.1| percent
identity: 61 in 429 aa"
/codon_start=1
/transl_table=11
/product="adenylosuccinate synthetase"
/protein_id="BAC19407.1"
/db_xref="GI:23494441"
/translation="MAAIVIGAQMGDEGKATDILGGLVDYVYVVKPNGNAGHTVV
VGKEYELKLLPAGLVBSTATPQNGVGVINLEALFEEDIGLEARGADARLRIRISANA
HLVAPYIMDRVORFLGKRAIGTTGRTGYODKYGRVGIQVQDIFDESILRQKI
ESALDVKQVLVKNRKAIVAEETVOYELSVADLRPMVIDATLELAKALDEGKHLV
MEGGQATMLMDVHGTYPPVTSNPTAGACVSGVGPTRIISTILGIIKAYTTRVAGNP
FPTLEFDKMGYLTQTVGGEVGVNTRKRCRWYDSVIARYASRVNGFTDYFLTKLDVL
TGIGSIPICVAYEVDVGRHDEMPMTQSDPFHAKPIFETMPAWDEDTGCRTFEELPQK
AQDYVRLLELSGARFSYIGVGPGRDQITVLHDVNES"
10063..10989
/notes="CE2598, similar to U23181-6|AAC48204.1| percent
identity: 20 in 247 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC19408.1"
/db_xref="GI:23494442"
/translation="MYAETLFTTTPRAEFHRNRTKVMHMLVLRCAAPSTDPFGA
PVAPVEHDLPEVPTRRDLMLDDAADVLPHDPTPSLDEIARQPDVAHLGTPQAPQH
PEQRIIRIVIGSDASLSAVITRLMRADNMMAEAVYVPTGESVAAKNWGLPADPGALR
LALTCAKVPVPTIRDDAAVAVAGSATITDWEPEITGEIIVDDHVLVRHEASPKTPRR
GIFGARLYPMVAVGIAAAMDTPAGTEIKKLFRRPTGTTLIPESLSGTGRATQAGGP
SLRVTVDGSRKRAVERVTFYRHLRLDLQVVRP"
11033..11203
/notes="CE2599"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC19409.1"
/db_xref="GI:23494443"
/translation="MTPAGPAQLLIILVVLVLFGSSKLPDVARSLGRSLRIFKSEIR
ELNRDDSTTNQP"

```

CDS complement(11200..12426)  
/note="CE2600, similar to X17313-3|CAA35192.1| percent  
identity: 74 in 359 aa"  
/codon\_start=1  
/transl\_table=11

Query Match 42.9%; Score 399; DB 1; Length 302070;  
Best Local Similarity 66.4%; Pred. No. 1e-84;  
Matches 592; Conservative 0; Mismatches 290; Indels 9; Gaps 1;

QY 27 CTGCGAATCGACGACCTACGAGCTTCATTTTCAGTCGCTCAATCAGGCCACCTCACCGA 86  
DB 194739 CATGAGATCGATGATCTCCGCTTTTCATCGAGGTTCGAGTCCGGGACCTGACGGA 194680

QY 87 ACTCGCGAAGATTAGGATCCCGAGCCGACATTTCCAGACCAATCAGCCGAGTGA 146  
DB 194679 GACGCGGATCGGCTCGGTATTGCGACGCCGACATGTCGCGCCGATGGCGCGTGA 194620

QY 147 AAAACACGCGACGCCACCTTTTCGACGCGCGCGCGCAAACTCGTCTCAACCAACG 206  
DB 194619 AAGACGATGGGTAGCGCGCTGTTGAGGCGCGCGCGGCGAGGCTCCAGCTCAACCCCG 194560

QY 207 AGGCAACGCTTCTCAACACGCGCGACGCTTCGTCGAGAAATCAACTCCGCGCAAC 266  
DB 194559 CGGGCAGCGCTTCTCCACACGCGCGCGCATACCGCGAGCTGGACATGGCGGTCA 194500

QY 267 TGAATCAACGCTCATGACCCGAGAAAGGCAATCCGACTGGACTTCATGCTATC 326  
DB 194499 GCAGTCAACGCTCGATGGACCCGAGCTGGGCGACGATCCGCTTGGATTTTCATG 194440

QY 327 CTTGGGCACTTGGATGGTCCCGAACTTATCCGAACATTCGCGCGCAACCCCAAGT 386  
DB 194439 GCTGGGCACTGGATGGTCCCGAGCTGATCCGTAATTCGCGACCGACCCCGCGGT 194380

QY 387 AGAATTCCAACTCCACGAGCGGCAATGCTCTCGTAGATCGTGTGCTGATGA 446  
DB 194379 GGAATTCGAGTGCACACGAGGCGCGCAATGAGTGTGTCGCGCTGTTGACGACG 194320

QY 447 RACTGACCTCGATTAGTTGGGCCCCAACCCTCGGAGTGTGTAATCTTTAGGTTGG 506  
DB 194319 CGCCGACCTCGCCCTCGTGGGACCGAGACCGCGAGGTGCGAGCTGTCTGGGTTGG 194260

QY 507 GCACTGCTTCTCAACGACTTGCCTAGCTGTTCCCGCAGATCACCGGCTTGCTCCTT 566  
DB 194259 GCGCTGCTGCGCAGGCTGCTGCTGCTGCGCGGCGGACCGCTCGAC----- 194205

QY 567 TTCTGCGCAAGAGAAATGCGCTTGAATTAATGCGCGCGAAGAACTTTGTCGCGATCG 626  
DB 194204 ----GGTGGAGCGCGTCAAGTCTGCGGAGGCTGCGGGAACCTTCGTGCGATGCT 194149

QY 627 AGCAGTTTCGCGACCGGCTCCTCATGATGCTAGCGGAGGCTGCGGAGGCTGCGG 686  
DB 194148 GCGCGTTTCGCGACCGGCTGCTTCTGACGCGCTGCGGAGGCTGCGGAGGCTGCGG 194089

QY 687 CAATGTGTTTCGAATCATGGAATCACCACCGCTCGCAGGCTTGTCAAGCGAGTCT 746  
DB 194088 CGACATGCTTTCGAATCATGGAATCACCACCGCTTCGCGCTGCTGCGGAGGCTCT 194029

QY 747 CGCGCTGTTGTTGTTGCGATGATGCTGATGCTTCCAGTGGGAATGTCGAACG 806  
DB 194028 GGGCGTGGGCTGCTGCGGCTGCGGAGGCTGCGGAGGCTGCGGAGGCTGCGGAGGCT 193969

QY 807 CCCACTTAGTCCACCGCTTATAGGAATAGGTTTGTGTTGTTGTTGTTGTTGTTGTT 866  
DB 193968 GCGCTTCGAACTGCGACCCACCGGAACTCGGCTGTTGTTGTTGTTGTTGTTGTT 193909

QY 867 GGCACCTCGGTGGATTAATCTCCGGAAGTTCTGTCGGGATCGAGGTATGC 917  
DB 193908 CGCCCGCGGCTGGAGGTGTTCCGGGATTTCTGTCGGCGGTTCCCGGTTTC 193858

RESULT 9  
AP006618\_15

WPCOMMENT  
Sequence split into 61 fragments LOCUS AP006618 Accession AP006618

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| AP006618_00   | 1       | 110000  |
| AP006618_01   | 100001  | 210000  |
| AP006618_02   | 200001  | 310000  |
| AP006618_03   | 300001  | 410000  |
| AP006618_04   | 400001  | 510000  |
| AP006618_05   | 500001  | 610000  |
| AP006618_06   | 600001  | 710000  |
| AP006618_07   | 700001  | 810000  |
| AP006618_08   | 800001  | 910000  |
| AP006618_09   | 900001  | 1010000 |
| AP006618_10   | 1000001 | 1110000 |
| AP006618_11   | 1100001 | 1210000 |
| AP006618_12   | 1200001 | 1310000 |
| AP006618_13   | 1300001 | 1410000 |
| AP006618_14   | 1400001 | 1510000 |
| AP006618_15   | 1500001 | 1610000 |
| AP006618_16   | 1600001 | 1710000 |
| AP006618_17   | 1700001 | 1810000 |
| AP006618_18   | 1800001 | 1910000 |
| AP006618_19   | 1900001 | 2010000 |
| AP006618_20   | 2000001 | 2110000 |
| AP006618_21   | 2100001 | 2210000 |
| AP006618_22   | 2200001 | 2310000 |
| AP006618_23   | 2300001 | 2410000 |
| AP006618_24   | 2400001 | 2510000 |
| AP006618_25   | 2500001 | 2610000 |
| AP006618_26   | 2600001 | 2710000 |
| AP006618_27   | 2700001 | 2810000 |
| AP006618_28   | 2800001 | 2910000 |
| AP006618_29   | 2900001 | 3010000 |
| AP006618_30   | 3000001 | 3110000 |
| AP006618_31   | 3100001 | 3210000 |
| AP006618_32   | 3200001 | 3310000 |
| AP006618_33   | 3300001 | 3410000 |
| AP006618_34   | 3400001 | 3510000 |
| AP006618_35   | 3500001 | 3610000 |
| AP006618_36   | 3600001 | 3710000 |
| AP006618_37   | 3700001 | 3810000 |
| AP006618_38   | 3800001 | 3910000 |
| AP006618_39   | 3900001 | 4010000 |
| AP006618_40   | 4000001 | 4110000 |
| AP006618_41   | 4100001 | 4210000 |
| AP006618_42   | 4200001 | 4310000 |
| AP006618_43   | 4300001 | 4410000 |
| AP006618_44   | 4400001 | 4510000 |
| AP006618_45   | 4500001 | 4610000 |
| AP006618_46   | 4600001 | 4710000 |
| AP006618_47   | 4700001 | 4810000 |
| AP006618_48   | 4800001 | 4910000 |
| AP006618_49   | 4900001 | 5010000 |
| AP006618_50   | 5000001 | 5110000 |
| AP006618_51   | 5100001 | 5210000 |
| AP006618_52   | 5200001 | 5310000 |
| AP006618_53   | 5300001 | 5410000 |
| AP006618_54   | 5400001 | 5510000 |
| AP006618_55   | 5500001 | 5610000 |
| AP006618_56   | 5600001 | 5710000 |
| AP006618_57   | 5700001 | 5810000 |
| AP006618_58   | 5800001 | 5910000 |
| AP006618_59   | 5900001 | 6010000 |
| AP006618_60   | 6000001 | 6021225 |

Continuation (16 of 61) of AP006618 from base 1500001 (AP006618 Nocardia farcinica IFM 1

Query Match 15.7%; Score 146; DB 1; Length 110000;  
Best Local Similarity 51.1%; Pred. No. 3.6e-24;  
Matches 403; Conservative 0; Mismatches 370; Indels 15; Gaps 2;

QY 15 CGACGGCGGAGACCTCGGATCGAGACCTAGGAGCTTCAATTCAGTCGCTCAATCAGG 74

DB 107923 CGGAGGTGGTGGCTGCTCGCGGAGGACCTGAGGTGGTACACGACACTCGTCGAGACACA 107982

```
QY 75 CCACCTCACGAAATGCCGAAAGATTAGGATATCCGAGCCACACATTTCCAGACGAAT 134
Db 107983 GAGCGTGAGCGCGCGCGCGGTGACCTCGCGAGCGACCTCTCGCGCATGCT 108042
QY 135 CAGCCGAGTGAATAACACGAGGACCCACATTTTCGACCGCGCGCGGCAAACTCGT 194
Db 108043 GCGCCGGCTCGAGCGGAGACTCGGGTGCCTTTTCGACCGCGCGGCAAACTCGC 108102
QY 195 CTTCAACACGAGGCGACGCTCTCTCAACACGCGAGCGCATCGTCGAGAAATCAA 254
Db 108103 GCTCAACGAGTTCGCGCGGTGTACTAGAACGCGCGCGGCGACAGTCCGAGCTCGA 108162
QY 255 CTCGCGCGCACTGAATCAAAACGCTCATGAGCCAGAAAAGGACCAATCCGACTGGA 314
Db 108163 CGCGCGCGACAGCGGTGCGGACCTGGCCGATCCGCGAGGGGTGTCGCGCTGTC 108222
QY 315 CTTTCATGATTCCTTGGGACATTTGGATGTTCGCCGAATTCGGAACATTCGCGCCGA 374
Db 108223 GTTCTGCAATTCCTTGGGCGGTGGTGTGTCGACAGCTGGTGGCGGGCTTCGCGCGG 108282
QY 375 ACACCCCAACGTAGAATTCACACTCCACCAAGCGGAGCAATGCTCTGCTAGATCGTGT 434
Db 108283 CGCGCGCGGTGAGGTGACACTGTGGAGGCGCGCGGCGGACCATCACCGCGCGGT 108342
QY 435 TTGTGCTGATGAATGACCTCGCATTAGTGTGSCCCCAAACTCGCGAGGTTGGTACCTC 494
Db 108343 GCTCGACGGAGCGCGACCTCGGCATGCTCTCGCGCGCGCGGCGGTGACGGCATC-- 108400
QY 495 TTGTAGGTGGCGGACCTGCTGCTCAACGATGTCGCGGCGCGGCGGCGGCGGCGGCGG 554
Db 108401 ---GGCTGGACCACTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 108456
QY 555 GCTTGCTCTCTTCTGCGCAAGAGAAATTCGCGTGTGATTACTGCGCGCGGAGAACTTT 614
Db 108457 GCTGGCC-----GGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 108507
QY 615 CGTGGCGATCGGAGCAGGTTTCGCGACCGGACTCTCTCATGATGATGATGATGATGATG 674
Db 108508 CGTCACGATGACCGCGGCTTGGGATGCGCGGCTCTTCGATGATGATGATGATGATG 108567
QY 675 CGGTTTGTTCGAATGCTGTTTCGAATCCATGGAATCATGAGGAGTTCGCGGCGGCTGT 734
Db 108568 GGGCATTCGCGCGGCTGATCGGTTGAGTCCAGTACCTGATCAGCGTTCGCGGCGGCTGT 108627
QY 735 CAGCGCAGTCTCGGCGTGTGCTGCTCGGATGATGATGATGATGATGATGATGATGATG 794
Db 108628 CGCGCGCGGCTCGGGTGCCTTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108687
QY 795 AATCGTC 802
Db 108688 GCGGTGC 108695
```

RESULT 10  
AP006618.16  
WPCOMMENT

Sequence split into 61 fragments LOCUS AP006618 Accession AP006618

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| AP006618_00   | 1       | 110000  |
| AP006618_01   | 100001  | 210000  |
| AP006618_02   | 200001  | 310000  |
| AP006618_03   | 300001  | 410000  |
| AP006618_04   | 400001  | 510000  |
| AP006618_05   | 500001  | 610000  |
| AP006618_06   | 600001  | 710000  |
| AP006618_07   | 700001  | 810000  |
| AP006618_08   | 800001  | 910000  |
| AP006618_09   | 900001  | 1010000 |
| AP006618_10   | 1000001 | 1110000 |
| AP006618_11   | 1100001 | 1210000 |
| AP006618_12   | 1200001 | 1310000 |
| AP006618_13   | 1300001 | 1410000 |

```
AP006618_14 1400001 1510000
AP006618_15 1500001 1610000
AP006618_16 1600001 1710000
AP006618_17 1700001 1810000
AP006618_18 1800001 1910000
AP006618_19 1900001 2010000
AP006618_20 2000001 2110000
AP006618_21 2100001 2210000
AP006618_22 2200001 2310000
AP006618_23 2300001 2410000
AP006618_24 2400001 2510000
AP006618_25 2500001 2610000
AP006618_26 2600001 2710000
AP006618_27 2700001 2810000
AP006618_28 2800001 2910000
AP006618_29 2900001 3010000
AP006618_30 3000001 3110000
AP006618_31 3100001 3210000
AP006618_32 3200001 3310000
AP006618_33 3300001 3410000
AP006618_34 3400001 3510000
AP006618_35 3500001 3610000
AP006618_36 3600001 3710000
AP006618_37 3700001 3810000
AP006618_38 3800001 3910000
AP006618_39 3900001 4010000
AP006618_40 4000001 4110000
AP006618_41 4100001 4210000
AP006618_42 4200001 4310000
AP006618_43 4300001 4410000
AP006618_44 4400001 4510000
AP006618_45 4500001 4610000
AP006618_46 4600001 4710000
AP006618_47 4700001 4810000
AP006618_48 4800001 4910000
AP006618_49 4900001 5010000
AP006618_50 5000001 5110000
AP006618_51 5100001 5210000
AP006618_52 5200001 5310000
AP006618_53 5300001 5410000
AP006618_54 5400001 5510000
AP006618_55 5500001 5610000
AP006618_56 5600001 5710000
AP006618_57 5700001 5810000
AP006618_58 5800001 5910000
AP006618_59 5900001 6010000
AP006618_60 6000001 6021225
```

Continuation (17 of 61) of AP006618 from base 1600001 (AP006618 Nocardia farcinica IFM 1

Query Match 15.7%; Score 146; DB 1; Length 110000;  
Best Local Similarity 51.1%; Pred. No. 3.6e-24;  
Matches 403; Conservative 0; Mismatches 370; Indels 15; Gaps 2;

```
QY 15 CGACGCGGAGACCTGCGAATCGAGACCTAGCAGCTTTCATGTCGCTCAATCAGG 74
Db 7923 CGAGGTGCTGCTCGCGAGACCTGAGTGTGTACAGACACTGTCGAGACACA 7982
QY 75 CCACCTCACGAACTGCCGAAAGATTAGGATCCGAGCCACACATTTCCAGACGAAT 134
Db 7983 GAGCGTGAGCGCGCGCGCGCTGACCTCGCGAGCGGACCTCTCGCGCATGCT 8042
QY 135 CAGCGGAGTGAATAACACGAGGACCCCACTTTTCGACCGCGCGCGGCAAACTCGT 194
Db 8043 GCGCGGCTCGAGCGCAGACTCGGGTGCCTTTTCGACCGCGCGGCAAACTCGC 8102
QY 195 CTTCAACCAACGAGGCGCAGCTTCTCTCAACGAGCGGCGCATGTCGCGAAGATTCAA 254
Db 8103 GCTCAACGAGTTTCGCGCGGTGTACTAGAAACACGCGCGCGCGCAGTCCGAGCTCGA 8162
QY 255 CTCGCGCGCACTGAATCAACGCGCTCATGGACCCAGAAAAGGACCAATCCGACTGGA 314
Db 8163 CGCGCGCGGACAGGCGGTGGCGGACCTGGCGCATCCCGCGAGGCGGTGGTGGCTGTC 8222
```

QY 315 CTTTCATGATTCCTTGGGCACTTGGATGTCCTCCGAACTTATCCGAAATTCCTGGCGCGA 374  
 DB 8223 GTTCTGCAATCTCTTGGGGGCTGGCTGGTGCCACAGCTGGTTCGGGGCTTCCGCGCGG 8282  
 QY 375 ACACCCCAACGTAGAAATTCACATCCACCAAGCGGAGCAATGCTCTCTGTAGATCGTGT 434  
 DB 8283 CGCCGCCCGGGTGAGGGTCACACTGTGGCAGGGCCCGCGCCACCATCACCCCGCGGT 8342  
 QY 435 TTTGGCTGATGAATGACCTGCACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 494  
 DB 8343 GCTGACGGGAGCGCGACCTCGGCATCGTCTCGCGCGCGCGCGCGCGCGCGCGCGCG 8400  
 QY 495 TTTAGGGTGGCGGCGCACTCTTCTGCAACGACTTTCCTCTAGCTGTTCGCGAGATCACCG 554  
 DB 8401 ----GGTGGAGCACTCTCTCGCGAGCGCGCTCATGTCTCGCGTTCGGCGCGCGAGCACCG 8456  
 QY 555 GCTTGCCTCTTTTCTGGCCAAAGGAGAAATGCGTTGATTAATCTGGCGGCGGAGAACCTTT 614  
 DB 8457 GCTGGCC-----GGCGCGCGCCAGGTGCGCTGGCGGAGCTGGCGAGCGCGAGTT 8507  
 QY 615 CTTGGCGATGCGAGCAGGTTCGGCACCGGACTCTCTCATGATGATTCATGCGGAGAACG 674  
 DB 8508 CGTACGATGCAACCGGGCTTCGGGATGCGCGCGCTCTTCGATGATCTGTGCGCGCGCGC 8567  
 QY 675 CGGTTTGTTCCTCAATGTGGTTTTCGAATCCATGAACTCACACCGCTCGCAGAGCTTGT 734  
 DB 8568 GGGCATCCCGCGGTATCGGTTTGAGTCCAGTCACTGATCATCGCTCGCGCGCTGGT 8627  
 QY 735 CAGCGAGCTCTCGCGGTGTGTGTTCGATGATGATGATGATGATGATGATGATGATGATG 794  
 DB 8628 CGCGCGCGGGCTCGGGTTCGCTGTGCGCGTGTGGACAGCGCGCCACACCGTGC 8687  
 QY 795 AATCGTGC 802  
 DB 8688 GCCGTGC 8695

RESULT 11  
 AF116906/c AF116906 7646 bp DNA linear BCT 26-APR-1999  
 LOCUS Rhodococcus equi virulence associated plasmid fragment d sequence.  
 DEFINITION AF116906  
 ACCESSION AF116906.1 GI:4680485  
 VERSION  
 KEYWORDS  
 SOURCE Rhodococcus equi (Corynebacterium equi)  
 ORGANISM Rhodococcus equi  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Nocardia; Rhodococcus.  
 Hines, S., Alperin, D., Prescott, J., and Nicholson, V.  
 Sequence of fragment d of 85 kb virulence associated plasmid of  
 Rhodococcus equi strain 103  
 Unpublished  
 JOURNAL  
 REFERENCE Hines, S., Alperin, D., Prescott, J., and Nicholson, V.  
 AUTHORS Direct Submission  
 TITLE Submitted (29-DEC-1998) Veterinary Microbiology and Pathology,  
 JOURNAL Washington State University, PO Box 647040, Pullman, WA 99164-7040,  
 USA

## FEATURES

Location/Qualifiers  
 1. .7646  
 /organism="Rhodococcus equi"  
 /mol\_type="genomic DNA"  
 /strain="103"  
 /db\_xref="taxon:43767"  
 /plasmid="virulence associated plasmid"  
 /note="fragment d"

## ORIGIN

Query Match 14.3%; Score 132.8; DB 1; Length 7646;  
 Best Local Similarity 51.1%; Pred. No. 5.4e-21;  
 Matches 373; Conservative 0; Mismatches 342; Indels 15; Gaps 2;

QY 35 TCGACGACCTAGCAGCTTTCATTTTCAGTCGCTCAATCAGGCGCCTCACCGAAATCTGCGG 94  
 DB 3952 TTGACGAACTCCGCTGGTTTGGGTTAGCAGAAATCGAGCACAATGCGCGGATGCGGGTGG 3893  
 QY 95 AAAGATTAGGATCCCGCAGCCCACTTTTCCAGACGAATCAGCGAGTGGAAACACG 154  
 DB 3892 ACGCCCTGCACATCTCTCAACCGACTCTCTCGCGGGGCTTGGCCGCAATTGAACGACAGG 3833  
 QY 155 CAGGACCCCACTTTTTCGACCGCGCGCGCGCAAACTCGTCTCTCAACCAACGAGGCGCAG 214  
 DB 3832 TTGGGGTGGCCCTGTTCGATCGGGTGAACCATCGCTGACGCTCAACGACTACGCGCGGA 3773  
 QY 215 CTTTCTCAACCAACCGCGCGCGCATTCGTCGAGAAATTCAACTCCGCGCAACTGAAATCA 274  
 DB 3772 TCATGTTGACATATCTCGCGGAGTTTGACCGAACTCCACTCAGCGCGCGAGGGATCG 3713  
 QY 275 AACGCTCATGAGACCCAGAAAGGCAAAATCCGACTGGAATTCATGATTCCTTTGGGCA 334  
 DB 3712 CTGCGCTCCGTAATCTCTGAGTCAGGACCGTTCGCTAGGCTTCTGCACTCCCTTGGCA 3653  
 QY 335 CTTGATGGTCCCGAACTTATCCGAACTTCGCGCGGCAACACCCCAACGTAAGATTC 394  
 DB 3652 GTTGGTTTGTGCTGAGATCATCCGGGTGTTTCGCGAGAGCGCGCGAGGCTGAACCTTCG 3593  
 QY 395 AACTCCCAAGCGCGCAGCAATGCTCTGCTGATGATCGTGTTCGCTGATGAACTGACC 454  
 DB 3592 CGCTGAACCGAGCTCCGAGCGAGGCAATCGAAGCGATGCTGCTCGGGGGAGATTGATG 3533  
 QY 455 TCGCATTAGTTGGCCCCCAACCTCCGAGGTGTTGTAATCTTTAGGGTGGCGCACTGCG 514  
 DB 3532 TCGCAGTGTATCGGGCCACCGCGCGCGG-----ACGAGTTGAGTGTATGAGCTCT 3479  
 QY 515 TTCGTCAACGACTTGGCTTCTGCTGAGATCGTGTTCGCGAGATCAGCGCTTGGCTCTTTTGGCC 574  
 DB 3478 ACGTAGAACGCTTGGCTTCTGCTGCGGAAATCTCACC-----CCTTAGCGGGC 3428  
 QY 575 AAGGAGAAATTCGCTTGAATTAATCTCGCGGAGAAACCTTTTCGTCGCGATGCGAGCAGTT 634  
 DB 3427 GTGCGACTATCCGATAGCAGACGCTTCGATAAGCTTTCATATGCTCGCAGGCCAT 3368  
 QY 635 TCGGACCCGACTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 694  
 DB 3367 TCGGGCTCCGCAAGCTGTGTCGAGAACTGTTTCGCGGCTGCTGCGGCTGTACCCCAATCG 3308  
 QY 695 TTTTCGATCCATGGAATCACCACCGTCGCGAGGCTGTGTCAGCGAGGCTCTCGGCGTTG 754  
 DB 3307 TCTTTGAGACCATGAGATCCCTACCTCTGAGGGCTGTGTCGCGAGGTTTCGGGGTGG 3248  
 QY 755 GTGTGGTTTC 764  
 DB 3247 CTGTGGTTTC 3238

## RESULT 12

AF116907  
 LOCUS Rhodococcus equi virulence plasmid, complete sequence.  
 DEFINITION AF116907 AF077324 AF077325 AF116908 AF118815 U60204  
 ACCESSION AF116907.2 GI:10657865  
 VERSION  
 KEYWORDS  
 SOURCE Rhodococcus equi (Corynebacterium equi)  
 ORGANISM Rhodococcus equi  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Nocardia; Rhodococcus.  
 Hines, S., Alperin, D., Prescott, J., and Nicholson, V.  
 Sequence of fragment d of 85 kb virulence plasmid of Rhodococcus equi  
 strain 103  
 Unpublished  
 JOURNAL  
 REFERENCE Hines, S., Alperin, D., Prescott, J., and Nicholson, V.  
 AUTHORS Direct Submission  
 TITLE Submitted (29-DEC-1998) Veterinary Microbiology and Pathology,  
 JOURNAL Washington State University, PO Box 647040, Pullman, WA 99164-7040,  
 USA

**TITLE** Development of a Rhodococcus equi-Escherichia coli plasmid shuttle vector  
**JOURNAL** Plasmid 38 (3), 180-187 (1997)  
**MEDLINE** 98096689  
**PUBMED** 9435020  
**REFERENCE** 3 (bases 1 to 80609)  
**AUTHORS** Takai, S., Hines, S.A., Sekizaki, T., Nicholson, V.M., Alperin, D.A., Osaki, M., Takamatsu, D., Nakamura, M., Suzuki, K., Ogino, N., Kakuda, T., Dan, H. and Prescott, J.F.  
**TITLE** DNA sequence and comparison of virulence plasmids from Rhodococcus equi ATCC 33701 and 103  
**JOURNAL** Infect. Immun. 68 (12), 6840-6847 (2000)  
**MEDLINE** 20536428  
**PUBMED** 11083803  
**REFERENCE** 4 (bases 68647 to 69906)  
**AUTHORS** Zheng, H. and Prescott, J.F.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (06-JUN-1996) Pathobiology, University of Guelph, Ontario, WA N1G 2W1, Canada  
**REFERENCE** 5 (bases 76635 to 80609; 1 to 1253; 18999 to 23927)  
**AUTHORS** Hines, S.A., Alperin, D.A., Prescott, J.F. and Nicholson, V.M.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (08-JUL-1998) Veterinary Microbiology and Pathology, Washington State University, P.O. Box 647040, Pullman, WA 99164-7040, USA  
**REFERENCE** 6 (bases 23922 to 25176)  
**AUTHORS** Hines, S.A., Alperin, D.A., Prescott, J.F. and Nicholson, V.M.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (29-DEC-1998) Veterinary Microbiology and Pathology, Washington State University, P.O. Box 647040, Pullman, WA 99164-7040, USA  
**REFERENCE** 7 (bases 69902 to 76640)  
**AUTHORS** Hines, S.A., Alperin, D.A., Prescott, J.F. and Nicholson, V.M.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (29-DEC-1998) Veterinary Microbiology and Pathology, Washington State University, P.O. Box 647040, Pullman, WA 99164-7040, USA  
**REFERENCE** 8 (bases 19006 to 19874)  
**AUTHORS** Byrne, B.A., Hines, S.A. and Prescott, J.F.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (11-JAN-1999) Veterinary Microbiology and Pathology, Washington State University, P.O. Box 647040, Pullman, WA 99164-7040, USA  
**REFERENCE** 9 (bases 68015 to 70595)  
**AUTHORS** Zheng, H. and Prescott, J.F.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (29-APR-1999) Pathobiology, University of Guelph, Ontario, WA N1G 2W1, Canada  
**REMARK** Sequence update by submitter  
**REFERENCE** 10 (bases 1 to 80609)  
**AUTHORS** Prescott, J.F., Hines, S.A., Nicholson, V.M. and Alperin, D.A.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (11-MAY-2000) Pathobiology, University of Guelph, Ontario, WA N1G 2W1, Canada  
**REMARK** Sequence update by submitter  
**COMMENT** On or before Oct 5, 2000 this sequence version replaced gi:4713919, gi:4680487, gi:4325102, gi:3834551, gi:3834552, gi:4680486.  
**FEATURES** Location/Qualifiers  
1..80609  
/organism="Rhodococcus equi"  
/mol\_type="genomic DNA"  
/strain="103"  
/db\_xref="taxon:43767"  
/plasmid="virulence plasmid"  
795..1154  
/note="ORF 1; part of pathogenicity island"  
/codon\_start=1  
/transl\_table=11  
/product="lysyl tRNA synthetase-like protein"  
/protein\_id="AAG21705.1"  
/db\_xref="GI:10657866"  
NAERFKDLEKWIAAAQVDRPNRSKSRMGSGSRTSKEGLAAIRNWAADNGYHVS

**gene** TRGRIAREIVEAFEAAH"  
**CDS** complement (1541..2059)  
/gene="vapG"  
complement (1541..2059)  
/gene="vapG"  
/note="part of pathogenicity island"  
/codon\_start=1  
/transl\_table=11  
/product="virulence associated protein VapG"  
/protein\_id="AAG21706.1"  
/db\_xref="GI:10657867"  
/translation="MSVRLTLAATLVAGISVLAPAGIANAEISMVSTTTAASVVEHAAN  
TYDFAEAKSGSSIPAKVAEQANSYSHGLVTLSTLAVYQHFSLTVGGGKTFDSDGGI  
SIPGVAVLEGTLFDLQHLXSDTVSFEYNAGVPLYNINFFDHSHTLLGHVQSGSIGT  
VSGIGGSGTGMQ"  
complement (3951..4706)  
/note="ORF 3; similar to Escherichia coli hypothetical  
protein encoded by GenBank Accession Number U82598"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAG21707.1"  
/db\_xref="GI:10657868"  
/translation="MNIRDNDPICABIDGGLYGOILSYCGSRISTWCNSLWRCPN  
SRIVKYNENIGARHLEYGARSANFLEQALSTATEIDILCKKQVLAHKHLTEQG  
LATRVICGVMAPIASMGYFDSIAASVFHRLPCTWSEKRLAIRHISHALKEDGVFF  
GATILGAGVRNLLGRTLCDIFNSRLGGFHNSCDDKAGLVDALAEASFEEVDVAVEGST  
AKFIARRPSEPDSPRLTISVGSDDLVCPOQL"  
4993..5829  
/note="ORF 4; part of pathogenicity island"  
/codon\_start=1  
/transl\_table=11  
/product="LyseR-like protein"  
/protein\_id="AAG21708.1"  
/db\_xref="GI:10657869"  
/translation="MRDAADALHISQPTLSRALARIERQVGVLPDRVNRHLQNLNDYG  
RIMLEHTRSLTELHSAGERIAALRNPSGTVRGLFHLASLWFPETIRVRESAPR  
VNFALNQAPSEAIAMLARGEIDVAIGVPPAADEFDWYELVERCLAVPESHPLAG  
RATIRIADASDNAFIMLRPFGLKLDLDELFAAGVVPQIVPETIEIPTLEGVAAAGF  
GVAVVPSRPRTKETEGVRYVPLDDVGAFRPGLAWPVGREPSVVTTRFLAFLNRQGE  
GA"  
6322..7230  
/note="ORF 5; part of pathogenicity island"  
/codon\_start=1  
/transl\_table=11  
/product="bicyclicmycin resistance protein-like protein"  
/protein\_id="AAG21709.1"  
/db\_xref="GI:10657870"  
/translation="MLAGVPVTAITYARELLPTKYTARAAGFVGGAAVGGWAGRLIS  
STLADSGWRAGLFASSGFSTAVVILFLVAPRRARHTRPPRAGRGVGMKSLVTGRQA  
AMYAQAPLLTGAFALYNNVGFRLSRDPNLSQQOIGLIYGTLAGTFAKSLAGCLVG  
AVGRPVLFGTGAIFLGGAVTLASNLWILVGLVLTAGFEAGHSVVGWGAESVA  
GTRASSIYTLAYGGASDFGLMGEAYOSYSPGLVAGVMFLTTIAAIFALLLTGER  
SAAEIVRQDMQVEAVQEGNDNAACP"  
7367..7930  
/gene="vapH"  
7367..7930  
/note="part of pathogenicity island"  
/codon\_start=1  
/transl\_table=11  
/product="virulence associated protein VapH"  
/protein\_id="AAG21710.1"  
/db\_xref="GI:10657871"  
/translation="MNLSTKTRKFLSRSTAVPATFMALTPWVGCAAPPLDPGPTHDL  
PTWREEGANYSDGTMNLVRASNPFLPSTHSDSGQQQWTVQVGLASALYVQRUKLVNVEG  
GETEGYAGLSFGGANWGTFTDNITQRLYRTDRTSEFNAVGPLYNVNFPDGHSAI  
LGHAQLGGVSSVIGIGGCTGTWIGDVA"  
7982..8419  
/note="ORF 7; part of pathogenicity island"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"

protein id="AAG21711.1"  
/db\_xref="GI:10657872"  
/translation="MQMSLKTIGLPLDLLSNPNRIONSPFPHPHVELTILVGLIIS  
RKGSKIVIPIDDELRKPNPAHGAFAVALEVGWIGDCFVLVPPVWDSADNQWKKVA  
DHVAGGNPAAGSKSFEDLVELTGACFGAVAHDKISPPPR"  
8534..9292  
/note="ORF 8; part of pathogenicity island"  
/codon\_start=1  
/transl\_table=11  
/product="two-component response regulator-like protein"  
/protein\_id="AAG21712.1"  
/db\_xref="GI:10657873"  
/translation="MESIRGISVILLADRPALTRATRVFELGMVVRAGNDVRLR  
SACAGFDLIVLDLPGMSGDELCKRDEQIFDPIPIILSSRNGEORTAGLESQA  
DYYILKPNVRELARIILDRIRREGNRPQRLPLSDEILQIYPSACIYNGD  
MITLSRFDLLAYFMKAPFFVHRSQDLLEWRKWSYGDLSITVIVHRRRLKRLGNYIS  
RIETVGRGVRYPGPVPPVQVAASDERLAGPRH"  
10278..10505  
/note="ORF 9; part of pathogenicity island"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAG21713.1"  
/db\_xref="GI:10657874"  
/translation="MRSGLPRLVHLVPRCSLAPSLALKRTGSEEVRIHMSDSSST  
SSIFHYLDLGSVPAVKIRRSFTPLRCMNR"  
10574..10759  
/note="ORF 10; part of pathogenicity island"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAG21714.1"  
/db\_xref="GI:10657875"  
/translation="MKYLPQHAHTSAIFNLIGRCRRSVPAFWARNLTAQERANIYAG  
HTPLAVIASLGGGTHF"  
10963..11325

Query Match  
Best Local Similarity 51.1%; Score 132.8; DB 1; Length 80609;  
Matches 373; Conservative 0; Mismatches 342; Indels 15; Gaps 2;  
QY 35 TCGACGACTACGCGAGCTTCATTTCAGTCGCTCAATCAGCGCACCTCACCGAACTGCGG 94  
DB 4949 TTGACGAACCTCGCTGGTGTGTGGCGTTAGCAGATCGGAGCACATCGCGATGCGGCTG 5008  
QY 95 AAAGATTAGGATCCGCGAGCCACATTCAGCAGGATCAGCGAGTGGNAACACAG 154  
DB 5009 ACGCCTTGACATCTCTCAACGACTCTCTCGCGGGCGCTTGCCCGCATTTGAACACAGG 5068  
QY 155 CAGGCACCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACGAGGCCACG 214  
DB 5069 TTGGGGTGCCCTGTTCGATCGGTGACCATCGCTGCGCTCAGCTACGCGCCGGA 5128  
QY 215 CTTCTCTCAACCAACGCGAGCCGATCGTGGAGAAATCAATTCACCGCGCAACTGAATCA 274  
DB 5129 TCATGCTTGAGCATACTCGCGGAGTTTGACCGAACTCCACTCAGCGCGGCGGATCG 5188  
QY 275 AACGCTCATGACCAAGAAAGGACACATCCGACTCGACTTCATGATCTCTTGGGCA 334  
DB 5189 CTGCGCTCGTAACTCTGAGTCAGGACCGTTTCGCTAGGCTTCCTGCACTCCCTTGGCA 5248  
QY 335 CTTGGATGGTCCCGAACTTATCCGAACATTCGCGCGCGCAACACCCCAACGTAGAATCC 394  
DB 5249 GTTGGTTGTGCTGAGATCATCCGGGTGTTTCGCGAGAGCGCGCGAGGTGAATTCG 5308  
QY 395 AACTCCACCAACGCGAGCAATGCTCTGTTGATAGTCGTGTTTGGCTGATGAATGAC 454  
DB 5309 CGCTGAACACGAGCTCGGAGGAGGCATCGAAGCGATGCTGCTCGGGGAGATTGATG 5368  
QY 455 TCGCATTAGTTGGCCCAACCTCGGAGTTGGTACCTCTTGGGTTGGGCGCCACCTGC 514  
DB 5369 TCGCAGTGTATCGGCGCACCGGCCCGCGCGG-----ACGAGTTCAGTGTATGAGCTCT 5422

QY 515 TTGCTCAACGACTTGCCCTAGCTGTCCGCGAGATCACCGGCTTGCCTCTTTCTGGCC 574  
DB 5423 ACGTAGAAGCGCTTTCCTCTGCTGTGCGCGAATCTCACC-----CCTTAGCGGCG 5473  
QY 575 AAGGAGAATTGCCGTGTGATTACTGTCGGCGGGAAGAACCTTTTCGTGCGGATGCGAGCAGTT 634  
DB 5474 GTGCGACTATCGGATAGCAGACGCTTCGATAAGCGCTTCATTATGCTCCGCGAGCCAT 5533  
QY 635 TCGGACCCGACTCTCTCATGATGATAGCGGAGAACCGCGTTTGTTCCTCATATGTTG 694  
DB 5534 TCGGCTCCGAAGCTGCTGACGAACTGTTTCGCGGCTGCTGCGCTGCTGACCCAAATCG 5593  
QY 695 TTTTCGAATCATGGAATCTCACCACTGCGCAGGCTTGTTCAGCGCAGGCTCTCGCGCTTG 754  
DB 5594 TCTTTGACGATCGAGATCCCTACCTCGAGGCGCTGTCGCGCGAGGCTTTCGCGGCTG 5653  
QY 755 GTGTGGTTCC 764  
DB 5654 CTGTGGTTCC 5663  
RESULT 13  
AP001204  
LOCUS  
DEFINITION  
AP001204  
ACCESSION  
AP001204.1 GI:10801055  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rhodococcus equi  
Rhodococcus equi  
Rhodococcus equi  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Nocardiaceae; Rhodococcus.  
REFERENCE  
AUTHORS  
Takai,S., Hines,S.A., Sekizaki,T., Nicholson,V.M., Alperin,D.A.,  
Osaki,M., Takamatsu,D., Nakamura,M., Suzuki,K., Ogino,N.,  
Kakuda,T., Dan,H. and Prescott,J.F.  
DNA sequence and comparison of virulence plasmids from Rhodococcus  
equi ATCC 33701 and 103  
Infect. Immun. 68 (12), 6840-6847 (2000)  
20536428  
PUBMED  
11083803  
REFERENCE  
AUTHORS  
Takai,S., Sekizaki,T., Kakuda,T., Nakamura,M., Suzuki,K. and  
Ogino,N.  
Direct Submission  
Submitted (18-FEB-2000) Shinji Takai, Kitasato University, School  
of Veterinary Medicine and Animal Sciences, Department of Animal  
Hygiene; Higashi 23-35-1, Towada, Aomori 034-8628, Japan  
(E-mail:takai@vmas.kitasato-u.ac.jp, Tel:81-176-23-4371(ex.459),  
Fax:81-176-23-8703)  
FEATURES  
Location/Qualifiers  
1..80610  
/organism="Rhodococcus equi"  
/mol\_type="genomic DNA"  
/strain="ATCC33701"  
/db\_xref="ATCC:33701"  
/db\_xref="taxon:43767"  
/plasmid="pREAT701 (p33701)"  
complement(join(80267..80610,1..184))  
/gene="orf64"  
complement(join(80267..80610,1..184))  
/gene="orf64"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB16673.1"  
/db\_xref="GI:10801119"  
/translation="MQGAGVLOQDFVDDSHRGFVDTFGALLIGIESLQGRGGS  
MRPPMAEPDALPCSSFTVRRVRAGGTARQRLACSLTVIVIASRMTSAARPWAA  
NRATCSLTARGSTCRGDSPLACGSLRGPWARSARFASAAAAPIRAVFEE  
SGRVSCSAQAASS"  
795..1154  
/gene="lsr2"  
gene  
CDS  
gene

|     |   |      |
|-----|---|------|
| CDS | 795..1154<br>/gene="lsr2"<br>/note="similar to Mycobacterium tuberculosis Lsr2 protein(112 aa) 1.9e-15(47.934% identity in 121 aa overlap) ; DAD:z95557-4"<br>/codon_start=1<br>/transl_table=11<br>/product="Lsr2 protein precursor"<br>/protein_id="BAB16610.1"<br>/db_xref="GI:10801056"<br>/translation="MARKVVTLTETFFDDFDGKIDDLSETIEFAVQGTSYRMDLRPK NAERPKOLEXWIAAAKVDGRPNRSKSRMGSGGSRTSKEGLAIRNAWDNGTHVS TRGRIAREIVEAFERAAH"<br>/complement(1541..2059)<br>/gene="vapG"<br>complement(1541..2059)<br>/gene="vapG"<br>complement(1541..2059)<br>/note="similar to Rhodococcus equi virulence-associated 15-17kDa antigen(189 aa) 4.7e-23(47.541% identity in 183 aa overlap) ; PIR:jc4072"<br>/codon_start=1<br>/transl_table=11<br>/product="R. equi VapG protein"<br>/protein_id="BAB16611.1"<br>/db_xref="GI:10801057"<br>/translation="MSVRTLLAATLVAGISVLAPAGIANAETSMTVTTAASSVEHAAN TYDFAEAGSSGIPAKVAEQAQNSYSHGLVTSLAYVQHFSLTVEGGGTFTGDSGGI SIPGVAVLGGTLFTEDLQHLYSDTVSFEYNAGVPYININFPDHSHTLLGHVQSGSIGT VSGIGGGTGGWQ"<br>complement(3951..4706)<br>/gene="orf3"<br>complement(3951..4706)<br>/gene="orf3"<br>/note="similar to Escherichia coli hypothetical protein(143 aa) 2.2e-12(42.478% identity in 113 aa overlap) ; DAD:u82598-30"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16612.1"<br>/db_xref="GI:10801058"<br>/translation="WNIRNDPICAIDGGLLYGQILSYCGSRISTWCNSLVWRCPN LATRVICGSMVAPIASPMGYPDSIAASFVHRLPGTWSEKRLAIRHISHALKEDGVFF GATILGAVRHNLCRTLCDIENSLGGFHNSCDDKAGLVDALEASFEEVDVAVEGST AKFIARRHSESPDESPTLTSVGSDDLVCQQL"<br>4993..5829<br>/gene="orf4"<br>4993..5829<br>/gene="orf4"<br>/note="similar to Bacillus subtilis LysR family of transcriptional regulator(278 aa) 3.8e-33(36.264% identity in 273 aa overlap) ; DAD:d26185-26"<br>/codon_start=1<br>/transl_table=11<br>/product="LysR family transcriptional regulator"<br>/protein_id="BAB16613.1"<br>/db_xref="GI:10801059"<br>/translation="MRDAADALHIQPTLSRALARIERQVGVPLFDRVNHRLQLNDYG RLMLEHRSRLTELHAGERIAALRNPESTGVRLGFLHSLASFWFPEIIRVPRESAPR VNFALNQASEIAEMLARGEIDVAVIGPPAAEDFDWYELVRLCLAVPESHPLAG RATIRIADASNAFTMLRRPFGRLKLVDELFAAGVVPQIVFETIEIPTLEGVLAAGF GVAVVPSRPRTKETGVRYVPLDDVCGAPRPIGLAWPVGPRESPVTVTRFLAFLANRQGE GA"<br>6322..7230<br>/gene="orf5"<br>6322..7230<br>/gene="orf5"<br>/note="similar to Escherichia coli Bicyclomycin resistance protein(417 aa) 2.6e-34(39.246% identity in 285 aa overlap) ; DAD:d90802-10"<br>/codon_start=1<br>/transl_table=11 | gene |
| CDS | 7367..7930<br>/gene="vaph"<br>7367..7930<br>/gene="vaph"<br>/note="similar to Rhodococcus equi virulence-associated 15-17kDa antigen(189 aa) 1e-23(44.086% identity in 186 aa overlap) ; DAD:d21236-1"<br>/codon_start=1<br>/transl_table=11<br>/product="R. equi Vaph protein"<br>/protein_id="BAB16615.1"<br>/db_xref="GI:10801061"<br>/translation="MNLSTTKRFLSRRTAVPATFVMAITVPMGCAAPPPLPDGPPTHDL PTMREEGANYSDDGTMLVPASSNFLEPSTHSDSGQOQTVQGVLASALVYQRLKLVNEG GETFEGYAGGLSFFGCGANVWGTLFTDNIQRLYDRTESFEFNAVGVLNVNFPDGHSAI LGHAQLGVSSVIGGGGTGWIGDVA"<br>7982..8419<br>/gene="orf7"<br>7982..8419<br>/gene="orf7"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16616.1"<br>/db_xref="GI:10801062"<br>/translation="MGMSLTKTIGLPLDLSNPRIQNSPPPPHRRVELTLVGLIDS RHGKSTKVIPIIDDELVRNPAHPGAPAALEVMGLGDCVFPLVPVWVSADNQWKKVA DHVAGGNFAAGSKKPFDELVBELTGACFYGAVAVHDKISPPFR"<br>8534..9292<br>/gene="tcr"<br>8534..9292<br>/gene="tcr"<br>/note="similar to Bacillus subtilis two-component regulator(240 aa) 3.8e-25(34.167% identity in 240 aa overlap) ; DAD:z99116-24"<br>/codon_start=1<br>/transl_table=11<br>/product="two-component response regulator"<br>/protein_id="BAB16617.1"<br>/db_xref="GI:10801063"<br>/translation="MESIRGISVLADRPALRTATRRVFEQLGMVVRAGNGDEVLR SACQGPDLIVLDVLPQMSGDELCKRDEQIFDPDIPILILSSRNGEQRVIAGLSGA DDYLLKPNRSRELALRIARILDRLEIRREGNRTPORLPLSDEILQLYPSARCIYVNGD MITLSRREDFLLAYFMKPPFVIHSRQDILLERWKWSYGDLSLTVTVHVRRLRKLGNYS RIETVWGRYWPVPPVPPVQAASDERLAGPRH"<br>10278..10505<br>/gene="orf9"<br>10278..10505<br>/gene="orf9"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16618.1"<br>/db_xref="GI:10801064"<br>/translation="NRSGLPRMLVHLVPRCSLAPSRALKRTGSEEVRIHMSDGSST SSIFHRYLDLGSVPVAVKIRRSFTPLRCMNR"<br>10574..10759<br>/gene="orf10"<br>10574..10759<br>/gene="orf10"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16619.1"  | gene |
| CDS | 795..1154<br>/gene="lsr2"<br>/note="similar to Mycobacterium tuberculosis Lsr2 protein(112 aa) 1.9e-15(47.934% identity in 121 aa overlap) ; DAD:z95557-4"<br>/codon_start=1<br>/transl_table=11<br>/product="Lsr2 protein precursor"<br>/protein_id="BAB16610.1"<br>/db_xref="GI:10801056"<br>/translation="MARKVVTLTETFFDDFDGKIDDLSETIEFAVQGTSYRMDLRPK NAERPKOLEXWIAAAKVDGRPNRSKSRMGSGGSRTSKEGLAIRNAWDNGTHVS TRGRIAREIVEAFERAAH"<br>/complement(1541..2059)<br>/gene="vapG"<br>complement(1541..2059)<br>/gene="vapG"<br>complement(1541..2059)<br>/note="similar to Rhodococcus equi virulence-associated 15-17kDa antigen(189 aa) 4.7e-23(47.541% identity in 183 aa overlap) ; PIR:jc4072"<br>/codon_start=1<br>/transl_table=11<br>/product="R. equi VapG protein"<br>/protein_id="BAB16611.1"<br>/db_xref="GI:10801057"<br>/translation="MSVRTLLAATLVAGISVLAPAGIANAETSMTVTTAASSVEHAAN TYDFAEAGSSGIPAKVAEQAQNSYSHGLVTSLAYVQHFSLTVEGGGTFTGDSGGI SIPGVAVLGGTLFTEDLQHLYSDTVSFEYNAGVPYININFPDHSHTLLGHVQSGSIGT VSGIGGGTGGWQ"<br>complement(3951..4706)<br>/gene="orf3"<br>complement(3951..4706)<br>/gene="orf3"<br>/note="similar to Escherichia coli hypothetical protein(143 aa) 2.2e-12(42.478% identity in 113 aa overlap) ; DAD:u82598-30"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16612.1"<br>/db_xref="GI:10801058"<br>/translation="WNIRNDPICAIDGGLLYGQILSYCGSRISTWCNSLVWRCPN LATRVICGSMVAPIASPMGYPDSIAASFVHRLPGTWSEKRLAIRHISHALKEDGVFF GATILGAVRHNLCRTLCDIENSLGGFHNSCDDKAGLVDALEASFEEVDVAVEGST AKFIARRHSESPDESPTLTSVGSDDLVCQQL"<br>4993..5829<br>/gene="orf4"<br>4993..5829<br>/gene="orf4"<br>/note="similar to Bacillus subtilis LysR family of transcriptional regulator(278 aa) 3.8e-33(36.264% identity in 273 aa overlap) ; DAD:d26185-26"<br>/codon_start=1<br>/transl_table=11<br>/product="LysR family transcriptional regulator"<br>/protein_id="BAB16613.1"<br>/db_xref="GI:10801059"<br>/translation="MRDAADALHIQPTLSRALARIERQVGVPLFDRVNHRLQLNDYG RLMLEHRSRLTELHAGERIAALRNPESTGVRLGFLHSLASFWFPEIIRVPRESAPR VNFALNQASEIAEMLARGEIDVAVIGPPAAEDFDWYELVRLCLAVPESHPLAG RATIRIADASNAFTMLRRPFGRLKLVDELFAAGVVPQIVFETIEIPTLEGVLAAGF GVAVVPSRPRTKETGVRYVPLDDVCGAPRPIGLAWPVGPRESPVTVTRFLAFLANRQGE GA"<br>6322..7230<br>/gene="orf5"<br>6322..7230<br>/gene="orf5"<br>/note="similar to Escherichia coli Bicyclomycin resistance protein(417 aa) 2.6e-34(39.246% identity in 285 aa overlap) ; DAD:d90802-10"<br>/codon_start=1<br>/transl_table=11 | gene |
| CDS | 7367..7930<br>/gene="vaph"<br>7367..7930<br>/gene="vaph"<br>/note="similar to Rhodococcus equi virulence-associated 15-17kDa antigen(189 aa) 1e-23(44.086% identity in 186 aa overlap) ; DAD:d21236-1"<br>/codon_start=1<br>/transl_table=11<br>/product="R. equi Vaph protein"<br>/protein_id="BAB16615.1"<br>/db_xref="GI:10801061"<br>/translation="MNLSTTKRFLSRRTAVPATFVMAITVPMGCAAPPPLPDGPPTHDL PTMREEGANYSDDGTMLVPASSNFLEPSTHSDSGQOQTVQGVLASALVYQRLKLVNEG GETFEGYAGGLSFFGCGANVWGTLFTDNIQRLYDRTESFEFNAVGVLNVNFPDGHSAI LGHAQLGVSSVIGGGGTGWIGDVA"<br>7982..8419<br>/gene="orf7"<br>7982..8419<br>/gene="orf7"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16616.1"<br>/db_xref="GI:10801062"<br>/translation="MGMSLTKTIGLPLDLSNPRIQNSPPPPHRRVELTLVGLIDS RHGKSTKVIPIIDDELVRNPAHPGAPAALEVMGLGDCVFPLVPVWVSADNQWKKVA DHVAGGNFAAGSKKPFDELVBELTGACFYGAVAVHDKISPPFR"<br>8534..9292<br>/gene="tcr"<br>8534..9292<br>/gene="tcr"<br>/note="similar to Bacillus subtilis two-component regulator(240 aa) 3.8e-25(34.167% identity in 240 aa overlap) ; DAD:z99116-24"<br>/codon_start=1<br>/transl_table=11<br>/product="two-component response regulator"<br>/protein_id="BAB16617.1"<br>/db_xref="GI:10801063"<br>/translation="MESIRGISVLADRPALRTATRRVFEQLGMVVRAGNGDEVLR SACQGPDLIVLDVLPQMSGDELCKRDEQIFDPDIPILILSSRNGEQRVIAGLSGA DDYLLKPNRSRELALRIARILDRLEIRREGNRTPORLPLSDEILQLYPSARCIYVNGD MITLSRREDFLLAYFMKPPFVIHSRQDILLERWKWSYGDLSLTVTVHVRRLRKLGNYS RIETVWGRYWPVPPVPPVQAASDERLAGPRH"<br>10278..10505<br>/gene="orf9"<br>10278..10505<br>/gene="orf9"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16618.1"<br>/db_xref="GI:10801064"<br>/translation="NRSGLPRMLVHLVPRCSLAPSRALKRTGSEEVRIHMSDGSST SSIFHRYLDLGSVPVAVKIRRSFTPLRCMNR"<br>10574..10759<br>/gene="orf10"<br>10574..10759<br>/gene="orf10"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16619.1"  | gene |
| CDS | 795..1154<br>/gene="lsr2"<br>/note="similar to Mycobacterium tuberculosis Lsr2 protein(112 aa) 1.9e-15(47.934% identity in 121 aa overlap) ; DAD:z95557-4"<br>/codon_start=1<br>/transl_table=11<br>/product="Lsr2 protein precursor"<br>/protein_id="BAB16610.1"<br>/db_xref="GI:10801056"<br>/translation="MARKVVTLTETFFDDFDGKIDDLSETIEFAVQGTSYRMDLRPK NAERPKOLEXWIAAAKVDGRPNRSKSRMGSGGSRTSKEGLAIRNAWDNGTHVS TRGRIAREIVEAFERAAH"<br>/complement(1541..2059)<br>/gene="vapG"<br>complement(1541..2059)<br>/gene="vapG"<br>complement(1541..2059)<br>/note="similar to Rhodococcus equi virulence-associated 15-17kDa antigen(189 aa) 4.7e-23(47.541% identity in 183 aa overlap) ; PIR:jc4072"<br>/codon_start=1<br>/transl_table=11<br>/product="R. equi VapG protein"<br>/protein_id="BAB16611.1"<br>/db_xref="GI:10801057"<br>/translation="MSVRTLLAATLVAGISVLAPAGIANAETSMTVTTAASSVEHAAN TYDFAEAGSSGIPAKVAEQAQNSYSHGLVTSLAYVQHFSLTVEGGGTFTGDSGGI SIPGVAVLGGTLFTEDLQHLYSDTVSFEYNAGVPYININFPDHSHTLLGHVQSGSIGT VSGIGGGTGGWQ"<br>complement(3951..4706)<br>/gene="orf3"<br>complement(3951..4706)<br>/gene="orf3"<br>/note="similar to Escherichia coli hypothetical protein(143 aa) 2.2e-12(42.478% identity in 113 aa overlap) ; DAD:u82598-30"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16612.1"<br>/db_xref="GI:10801058"<br>/translation="WNIRNDPICAIDGGLLYGQILSYCGSRISTWCNSLVWRCPN LATRVICGSMVAPIASPMGYPDSIAASFVHRLPGTWSEKRLAIRHISHALKEDGVFF GATILGAVRHNLCRTLCDIENSLGGFHNSCDDKAGLVDALEASFEEVDVAVEGST AKFIARRHSESPDESPTLTSVGSDDLVCQQL"<br>4993..5829<br>/gene="orf4"<br>4993..5829<br>/gene="orf4"<br>/note="similar to Bacillus subtilis LysR family of transcriptional regulator(278 aa) 3.8e-33(36.264% identity in 273 aa overlap) ; DAD:d26185-26"<br>/codon_start=1<br>/transl_table=11<br>/product="LysR family transcriptional regulator"<br>/protein_id="BAB16613.1"<br>/db_xref="GI:10801059"<br>/translation="MRDAADALHIQPTLSRALARIERQVGVPLFDRVNHRLQLNDYG RLMLEHRSRLTELHAGERIAALRNPESTGVRLGFLHSLASFWFPEIIRVPRESAPR VNFALNQASEIAEMLARGEIDVAVIGPPAAEDFDWYELVRLCLAVPESHPLAG RATIRIADASNAFTMLRRPFGRLKLVDELFAAGVVPQIVFETIEIPTLEGVLAAGF GVAVVPSRPRTKETGVRYVPLDDVCGAPRPIGLAWPVGPRESPVTVTRFLAFLANRQGE GA"<br>6322..7230<br>/gene="orf5"<br>6322..7230<br>/gene="orf5"<br>/note="similar to Escherichia coli Bicyclomycin resistance protein(417 aa) 2.6e-34(39.246% identity in 285 aa overlap) ; DAD:d90802-10"<br>/codon_start=1<br>/transl_table=11 | gene |
| CDS | 7367..7930<br>/gene="vaph"<br>7367..7930<br>/gene="vaph"<br>/note="similar to Rhodococcus equi virulence-associated 15-17kDa antigen(189 aa) 1e-23(44.086% identity in 186 aa overlap) ; DAD:d21236-1"<br>/codon_start=1<br>/transl_table=11<br>/product="R. equi Vaph protein"<br>/protein_id="BAB16615.1"<br>/db_xref="GI:10801061"<br>/translation="MNLSTTKRFLSRRTAVPATFVMAITVPMGCAAPPPLPDGPPTHDL PTMREEGANYSDDGTMLVPASSNFLEPSTHSDSGQOQTVQGVLASALVYQRLKLVNEG GETFEGYAGGLSFFGCGANVWGTLFTDNIQRLYDRTESFEFNAVGVLNVNFPDGHSAI LGHAQLGVSSVIGGGGTGWIGDVA"<br>7982..8419<br>/gene="orf7"<br>7982..8419<br>/gene="orf7"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16616.1"<br>/db_xref="GI:10801062"<br>/translation="MGMSLTKTIGLPLDLSNPRIQNSPPPPHRRVELTLVGLIDS RHGKSTKVIPIIDDELVRNPAHPGAPAALEVMGLGDCVFPLVPVWVSADNQWKKVA DHVAGGNFAAGSKKPFDELVBELTGACFYGAVAVHDKISPPFR"<br>8534..9292<br>/gene="tcr"<br>8534..9292<br>/gene="tcr"<br>/note="similar to Bacillus subtilis two-component regulator(240 aa) 3.8e-25(34.167% identity in 240 aa overlap) ; DAD:z99116-24"<br>/codon_start=1<br>/transl_table=11<br>/product="two-component response regulator"<br>/protein_id="BAB16617.1"<br>/db_xref="GI:10801063"<br>/translation="MESIRGISVLADRPALRTATRRVFEQLGMVVRAGNGDEVLR SACQGPDLIVLDVLPQMSGDELCKRDEQIFDPDIPILILSSRNGEQRVIAGLSGA DDYLLKPNRSRELALRIARILDRLEIRREGNRTPORLPLSDEILQLYPSARCIYVNGD MITLSRREDFLLAYFMKPPFVIHSRQDILLERWKWSYGDLSLTVTVHVRRLRKLGNYS RIETVWGRYWPVPPVPPVQAASDERLAGPRH"<br>10278..10505<br>/gene="orf9"<br>10278..10505<br>/gene="orf9"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16618.1"<br>/db_xref="GI:10801064"<br>/translation="NRSGLPRMLVHLVPRCSLAPSRALKRTGSEEVRIHMSDGSST SSIFHRYLDLGSVPVAVKIRRSFTPLRCMNR"<br>10574..10759<br>/gene="orf10"<br>10574..10759<br>/gene="orf10"<br>/codon_start=1<br>/transl_table=11<br>/product="hypothetical protein"<br>/protein_id="BAB16619.1"  | gene |





```

GSAESEA VDP EIAAARQLEAGDFEAARTSYQAILDANPASTEAKGAIRQIDFLTRAT
AQRPAVADAPAGDIDAAFAAADVQIQVDQGAAPFRLIALVENTSGDDRVRVTRR
I VLEPDLFPADPEVAGGRNLIANALY"
/ complement (5652..6821)
/ gene="fadA4"
/ locus_tag="MAP2436c"
/ product="fadA4"
/ complement (5652..6821)
/ gene="fadA4"
/ locus_tag="MAP2436c"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAS04753.1"
/ db_xref="GI:41396892"
/ translation="MIVAGARTPIGKLMGSLKDFSGADLGAIAIAGALEKADVPADLV
EYVINGOVLTCAGOMPARQAAVAGKIGDWDVPALSIKMKCLSGIDAIALAOLIRAGE
FVIVAGQESQTKAPHLLMDSRAGYGDVTYLVHMYADGLHDVDFPTQPGALTEQR
NDVFTREEQDYAARSHQRAAAAKWQGFDEVPVNI PORKQDPLQFTDEDEGIRA
NTTASLAGLKPAFRKDTGITAGSASQISDGAAGVVMNSKAKAQQGLSLWABIGAHG
VVAGPDSLQSQPANAIKKALREGISVDQLDVVEINEAFVAVALASTRELGVNPDIV
NVNGAIAVGHPIGMSGARI TLHAALELSRRSGSYVAVALCGAGGQGDALILRAG"
6915..7379
/ locus_tag="MAP2437"
6915..7379
/ locus_tag="MAP2437"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAS04754.1"
/ db_xref="GI:41396893"
/ translation="WTTQVDARQVLATVTAIDHVGIVADLDAAIAWYHDLGMI
LVHEVNEEQIGREMLAVRGAPAGTAQLMAPIDDSSTIAKFLDKRGPGIQQMAMR
VSDLTTLERLEQGVRLIYDAPRGRGTANSRINFTHPKDAGGVLIELVEPAA"
/ complement (7390..7686)
/ locus_tag="MAP2438c"
/ complement (7390..7686)
/ locus_tag="MAP2438c"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAS04755.1"
/ db_xref="GI:41396894"
/ translation="WARRTPPRRQLLGGPPALRRVEVPGDHDYEVYRPIAARAVK
TVRCPCDHEIRSGAAHLVWVPVPTLGGAEERHWHTCWAHRATRGPTRRWS"
/ complement (7710..8390)
/ locus_tag="MAP2439c"
/ complement (7710..8390)
/ locus_tag="MAP2439c"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAS04756.1"
/ db_xref="GI:41396895"
/ translation="MSRVLVIAQCTVDYVGRLTALFSLPARRLLLFKADGSYSVHADH
RAYKLNMSPKCLMREAGADAAVVMVENKAGEQLRITVEDIEHDSHDLGVDPLV
KQGVSAHLQALLAEHVOLIGEVTLVRREYMTAIGPVDLLCRDERGSGSVAVEIKRGE
IDGVQLTRYLLNDRDSVLAPVRGVFAAQIIPQARTILATDRGIRCTVLDYDKRGM
DSDYKRLF"
8427..10061
/ locus_tag="MAP2440"
8427..10061
/ locus_tag="MAP2440"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAS04757.1"
/ db_xref="GI:41396896"
/ translation="MSATOSTAQIRGVLEKTIROSGRIPETPAYGSLILGRVTSOH
RRIRIQIIMTVWLGANLIGIGALILVIVAI PAPSVPNDAPANIITGVGPAYVASA
LAVGVTIITRVLARWAIERKPTPTDNRNIFLAPMRIAULVDLVLMAGTLLTLVL
YGLVDTMFIPRFYFGVFCGVLVATACYLLAEFALRPVAAQALEAGPPPRRLTAGIMG
RTWTVFVFGSGVYIGIALALLFEIWLRLNLTETEFAGVLIIVSTAALIFGCILLMILA
WLTAIPVRVVRVAAALKRVERGDLRGDLVVPDGTGELGELQRFNMGVNDGLRERVRDLF

```

```

GRHVRVVALAERERPKLGEERHVA VVFDIVGSTOLVTRRPPAEVSVVLNQFFGI
VVEEDRCHGVNKFEGDATALIFGAPNLDPCPEDAALTAARAIALDRLANEMPQCRAG
IGVAGQVAGNVGARBERFEYTVIGEPVNEARCELAKSHSRLLATGDAIEGASEK
BRARSLGDTVTTLRGHERPTRLASPAGAADRPAGTD"
/ complement (10525..11787)
/ locus_tag="MAP2441c"
/ complement (10525..11787)
/ locus_tag="MAP2441c"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAS04758.1"
/ db_xref="GI:41396897"
/ translation="MPARSVMVMPHGPVHAQAGTRGYRMTAALCGAGLASFAMMYCS
QALLPALSAHRYRIGPATAAALVTSLTTGALALSII PASVLSRYGRIRVWLISGVASSV
IGLLPSPSIVGLLVFGRAAQGVALAGIPAVAMALLAEVDASSIGSAMGRVIAGTTI
GLAGRIVPSVVQVGTWRVALLACSLITLAGTAVFVAVLVPSSREFTPKPASVRAALR
NLGHRNRPVIAKLPAVGFLVLMGGFVTVYNTLGRVLAARPFGLAPSVVGLLFLLLVVG
TGTSVAGRLADRRERPLVGAALPIAVAGLLLTVPATLAAI VAGVGVTGTGFFAAHT
VASGVGVAQDRRAEASALYFSYTLGSGSVAGAFGGVLYGVGGWSATVCFVYVLLMA
GAALVALLVRONGFRIGRRVVTVSASVK"
11854..12738
/ locus_tag="MAP2442"
11854..12738
/ locus_tag="MAP2442"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAS04759.1"
/ db_xref="GI:41396898"
/ translation="MAVDGYEFHTTIAELOHVTAAAEQLHVAQPTLTMLRLAEORLG
VALPDRHRLSLNTYGRIFRYEHARRAOLEDSARRAIALADPAVGSIRLAYLSFG
STVPRLLIAREKSSPRVTFLKGAESIADRVRSGGVDIGVSVPRPVERTLAWRSL
FRQRLGVAVPDGHREARCAVSMVDLADEPFTVTHPGFGMRLLDELCAAQFRPRVA
LQAPNLTTAASLIIVAAGLISLVPIIDGSSYPGVSVKPLADADADRVDGKHWDSGRPLP
RSARDFTAAAAAAGVGGW"
12825..14348
/ gene="alkA"
/ locus_tag="MAP2443"
12825..14348
/ gene="alkA"
/ locus_tag="MAP2443"
/ codon_start=1
/ transl_table=11
/ product="AlkA"

Query Match 14.0%; Score 130.2; DB 1; Length 302325;
Best Local Similarity 48.5%; Pred. No. 2.1e-20;
Matches 431; Conservative 0; Mismatches 443; Indels 15; Gaps 2;

QY 16 GACGGCGGAGACCTGCGAATCGACGACCTACGACGCTTCATTTCAGTCGCTCAATCAGGC 75
Db 11842 GAGTGGCGGGATGGCAGTGGTGACTACGAGTGTTCATCAGTGGTTCATCAGCTCGCGAGCTGCAG 11901

QY 76 CACCTCACCGAAACTCGCGAAAGATTAGGCATCCGCGAGCCACACTTTTCAGAGCAATC 135
Db 11902 CAGCTCACCGCGCGCGCGAGCTGCACGTCGACGCGACGCTGTACCCGGAGTCTG 11961

QY 136 AGCCGAGTGGAAAAACACGAGGCAACCCCACTTTTCGACCGCGCGCGCGCAAACTCGTC 195
Db 11962 GCGCGCTTGGAAACAGCGACTCGCGCGTGGCGTGTTCGATCGTCACGCGCGGAGCTGTC 12021

QY 196 CTCACCAACGAGGCGCAGCCCTTCCTCAACACGCGAGCGCCATCGTCGAGAAATTCAC 255
Db 12022 CTCACACTGAGCGCGCATCTTCTACGAAACGCGCGCGCGCGGCGAGTTGGAGCTGGAT 12081

QY 256 TCGCGCGCAACTGMAATCAACCGCTCATGGACCCAGAAAAAGGCACCAATCCGACTGGAC 315
Db 12082 TCGGCACGGGTGCGATAGCGGATCTGCGGATCCCGCGCTCGCGAGATCCGGTTGGCC 12141

QY 316 TTCATGCATTCCTTGGGCATCTGGATGGTCCCCGAACTTATCCGAACATTTCGCGCGGAA 375
Db 12142 TACCTGAGTTCATTCCGATCAACCGTGGTGGCGCGTGTGATCGCCCGGTTCAGGAGTCG 12201

```

|            |       |  |   |
|------------|-------|--|---|
| QY         | 376   | CACCCCAACGTAGAAATTTCCAACTCCACCAAGCGGAGCAATGCTCCTCGTAGATCGTGT   | 435   |
| Db         | 12202 | TCCTCCGGGTGACCTTCAGCTTGAGAAAGCGCGCGAATCATGCTCCGATCGCGTG        | 12261   |
| QY         | 436   | TTGGCTGATGAACACTGACCTCGCATTTAGTTGGCCCCCAACCTGCGGAGTTTGGTACCTCT | 495   |
| Db         | 12262 | CGGTCCGGCGGTGACATCGGGTGGTATCCCCCGACCGTTCGAG-----CGAAC          | 12315   |
| QY         | 496   | TTAGGGTGGCGCCACTGTTCTGTCAAACACTTGGCTTACCTAGCTTGTCCCGCAGATCACGG | 555   |
| Db         | 12316 | CTGGCTGGCGCAGCGCTGTTTCGCCAGCGCCTCGGTGTCGGGTGCCGATGTCATCGA      | 12375   |
| QY         | 556   | CTTGCTCTCTTTCTTGGCCAAAGAGAAATTGCCGTGATTACTGCGCGCGAAGAACTTTC    | 615   |
| Db         | 12376 | TTCCGC-----CGCGTCCGGGTGTCATGTTGGACTGCGCGACGACCTTTC             | 12426   |
| QY         | 616   | GTGGCGATCGAGCAGGTTTCGGCACCGACTCCTCATGATGATAGCCGAGAACGCC        | 675   |
| Db         | 12427 | GTGCGATGCACCCCGTTTCGGATGCGAGCGCTCCTCGACGAGCTTTCGCGCCAGCC       | 12486   |
| QY         | 676   | GTTTTGTTCCCAATGTGTTTTGCAATCCATGGAATCACACCGTTCGAGGGTTGTC        | 735   |
| Db         | 12487 | CAATTCCCGCCTCGGTGTCCTGCAAGCGCGGAATCTCACGACCGCGCGCAGCTAGTC      | 12546   |
| QY         | 736   | AGCGCAGGTCTCGCGTGTGTTGTTGTTCCGATGATGATCCGTACTTCCACAGTGGGA      | 795   |
| Db         | 12547 | GCGCGGGCTGGATCAGCTGTCCGATCGACGCGAGCATTCCTCGCGGAGTA             | 12606   |
| QY         | 796   | ATCGTGCAACGCCCACTTAGTCCACCCCGCTTATAGGAACTAGTTTGGTGTGGCACTC     | 855   |
| Db         | 12607 | TCCGTCAAGCGCTGCGCGACGCGATGCTACCGCGATGTCGGCATGATCTGGATTC        | 12666   |
| QY         | 856   | AACGCGGGCGGCACTCGGTGATACCTTCCGGAAGTTCTGTCGGCG                  | 904   |
| Db         | 12667 | GCCCGCGCTTCCCGGTTCCGCGCGCACTTCATCGCGCGCGCGCG                   | 12715   |
| RESULT 15  |       |  |   |
| AP235504/c |       | 77534 bp   | DNA linear BCT 19-JUL-2000  |
| LOCUS      |       |  |   |
| DEFINITION |       |  | Streptomyces hygroscopicus var. ascomyceticus PK520 biosynthetic gene cluster, partial sequence.  |
| ACCESSION  |       |  | AP235504  |
| VERSION    |       |  | AP235504.1 GI:9280381   |
| KEYWORDS   |       |  | Streptomyces hygroscopicus subsp. ascomyceticus   |
| SOURCE     |       |  | Streptomyces hygroscopicus subsp. ascomyceticus   |
| ORGANISM   |       |  | Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  |
| REFERENCE  |       |  | 1 (bases 1 to 77534)  |
| AUTHORS    |       |  | Wu, K., Chung, L., Revill, W.P., Katz, L. and Reeves, C.D.  |
| TITLE      |       |  | The PK520 gene cluster of Streptomyces hygroscopicus var. ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units |
| JOURNAL    |       |  | Gene 251 (1), 81-90 (2000)  |
| MEDLINE    |       |  | 20323220  |
| PUBMED     |       |  | 10863099  |
| REFERENCE  |       |  | 2 (bases 1 to 77534)  |
| AUTHORS    |       |  | Reeves, C.D., Wu, K., Chung, L., Revill, P. and Katz, L.  |
| TITLE      |       |  | Direct Submission   |
| JOURNAL    |       |  | Submitted (16-FEB-2000) Korean Biosciences Inc, 3832 Bay Center Place, Hayward, CA 94545 USA  |
| FEATURES   |       |  | Location/Qualifiers   |
| source     |       |  | 1..77534  |
|            |       |  | /organism="Streptomyces hygroscopicus subsp. ascomyceticus"   |
|            |       |  | /mol_type="genomic DNA"   |
|            |       |  | /variety="ascomyceticus"  |
|            |       |  | /db_xref="taxon:132248"   |
| gene       |       |  | complement(412..1836)   |
|            |       |  | /gene="fkbw"  |
| CDS        |       |  | complement(412..1836)   |

```
/gene="fkbw"
/codon_start=1
/transl_table=11
/product="fkbw"
/protein_id="AAF86380.1"
/db_xref="GI:9280382"
/transl_table="MVLPLGVGRPRRTGLALTAAVAVLLPSPATARPAPDAPDG
HCAELSQLRVPAARQAQDCLELTTAGTAATGHTDADWAGLTPKNLATPTGVPGLQ
IDGFPDTSINTNHNHDAQFVIRLPDRNGLVAGTFCNREQYANDRAIADWL
LVTGMSGGYLRWOLENHPGLYDGVWEGTLMRAECPNLFTLEPALRDYPTAAG
GPAAGQAREALLRAGFPADSEFLWPYHVEHVTDLTORVREEFPDGFGEKEAGTVC
APGTPGCDADYDVATSSGVAHAVERIALTGRIGKPLITTIQALSTCCCLPPGLRRVR
AYVQRPGRGVRYRIEDGTHTDALVDTFPDLRLPLPCHRSATLESMLTTGHRPP
ANHVTVPAGADQASLLAECRLDG"
complement(2020..3579)
/gene="fkbw"
/gene="fkbw"
/codon_start=1
/transl_table=11
/product="polyhydroxybutyrate depolymerase"
/protein_id="AAF86381.1"
/db_xref="GI:9280383"
/transl_table="MQPPFPGILTLPLFLSSPPVSLSRPGRGVLTRLVAVVALV
LGALLGPATAAAGLAKGLTKADLTVADFGTNPRLNMTVYRPAISPAPAVV
EALHCTQDAGQADNSGLLSFADRYGFLVFAETTSNNANRCFNWFQSSDNRGQG
FAASIRQAAHTVSAYGADPQRTYITGLSAGGAMTSVMLATYPDVFOGAVVAGLPFG
CARTDVSSAYLCMNPPTDLTADQWARRVDRDGYPSWSGPWRVAVIWHGDKDTTVAERNAD
ELRDQWTAHVGSQTPDRTSVIGNSRHEEYLAADGSAVAVNVRVPGIGHGTVPDPC
TGACQCGSTGAAVFLDSCSYWITQFFGLSGASDPGSLPAPSLAATGATDTTISL
TWKRDGATDYAVHRDGAQITTSATTSYDTGLRAGTSHTYAVARADADCKAGPLSG
VTAQGTATAVCWTAAGYAHVQAGRATTSAGTYAKSGQNMGLYNTFVTTLKESPT
GYFTVANDTCP"
complement(3969..4496)
/gene="fkbw"
/gene="fkbw"
/codon_start=1
/transl_table=11
/product="transcription regulator"
/protein_id="AAF86382.1"
/db_xref="GI:9280384"
/transl_table="MVEGWQRYPEDFTTGCVYRHPGRVLETDNSWLTLLTONTAP
LHEDRHSAGTTWCKPLVDSTFTLALVTGQSVDTISOHVMANLGNDRVRLPNPFEGD
TIYSQSEVLNTRRESASRPDVGIIISVRTIGYNQDGVITFTERTLMVYRRGGRFAAV
EPWWDERSRRAAGLH"
complement(4595..5488)
/gene="fkbw"
/gene="fkbw"
/codon_start=1
/transl_table=11
/product="transcription regulator"
/protein_id="AAF86383.1"
/db_xref="GI:9280385"
/transl_table="MEIRELQWFTALAEHEHVTITABRLNISQTLGRAIRRLRRVVG
VPLFDRHQRNRLNLYKGEVFRALRAISEVSGAEQRTITALIDPDTGTVALGFLHSYG
TWLVPDLGAVYRLAPGTTFELRGDAADAVDDVDRNSRLDLGLTSPPRAGDDVWVTP
RDEALCLLVPPGHRLARRRVRTAELADEAFVALPVGRLQITDRLCAAAGFVRV
LESTELSTRLVAGLGVAVAPLTGSPGFGDHAVAPLDDPQARRTVGVVTLGAGP
RAPVVARFYDVRRAHAPHV"
5601..6818
/gene="fkbw"
5601..6818
/gene="fkbw"
/note="L-carnitine dehydratase"
/codon_start=1
/transl_table=11
/product="fkbw"
/protein_id="AAF86384.1"
```

```

/db_xref="GI:9280386"
/translation="MSEPHRPEQERPAGEPLSGLLVLSLEQAAVAFPAFTRHLADLGR
VIKIERPGSGDLARGDRYTRVGMSSHFVWLNKESQVOLDVRSPEKRNHLHALVDRAG
VLVQNLAAGAAGLASATRSSGATEADHLRTVPATAPAAATADKAYDOLLVQCERGL
VSTGTPEPSPKVLGIADI CAGMAYSGILITALLKRAETGRGSOLEVSMALGEMW
GYAEYTRGCTAPARAGASHATIAIYGFPTTRDGGTINLGLQNEKRWASFCQVLIQR
PGLCDPRSGNADRVAHRTEDALVSEVITGLTGEELVAREEASIAIYARQVTRFEE
SEHPQDRGRWAPDSPVGALEGLIPVTPHGEHPRRLGRVPELGEHTESVLAWLAA
PHSADREAGHAE"
gene
6808..8052
/genes="fkbF"
CDS
6808..8052
/genes="fkbF"
/notes="fkbF"
/notes="similar to Rhizobium trifolii MatC; dicarboxylate
carrier"
/codon_start=1
/transl_table=11
/product="fkbF"
/protein_id="AAF86385.1"
/db_xref="GI:9280387"
/translation="MENELTGVLIILAAVFLLAGVRGLNMGILLALVATFLLGVVALDRT
PDEVLAGFASPLVAVTFELFGIARVNGTVDLVRVAVRAGVAVPWLGLLA
ALLCARGASPAVAIVAPISVAFVRHRIIDPLIYAGLMVANGAAGSAPSGILGGIV
HGALEKHLIPVSGLLPAGTIPFNLAVAVSWLVGLRRLEHDHDEDDTDEGDPAS
RGAERHMTLTAAALVLTGTVLSLDTGLFALTALALLPFTSQQATKSIAPVW
LIVCGITVVALIQELGIVDSLSGMIIAITGLLALALVICYGVGVVSAFASPTGILGA
LNPISBPLKSGAIGTGMVMAAAATVVDASPFETNGALVAVANAPERLPGVYQGL
LWAGVGCALAPAAWAAFWA"
gene
8156..8824
/genes="fkbG"
CDS
8156..8824
/genes="fkbG"
/notes="O-methyl transferase"
/codon_start=1
/transl_table=11
/product="fkbG"
/protein_id="AAF86386.1"
/db_xref="GI:9280388"
/translation="MANQITLSDTLAYVRKSLRDEVLRLAQTAELPGGVLPV
QAEQGFLEFLRLTGARQVLEIGTVGYSTLCIARGAPGGVTVCDVMPKPEVGE
RWEEAGVADRIVRGIDARTVLITGLLDEAGAPSPDMWFIDADKAGPAYEAAALP
LVRRGLIVVDNTLFFGRVDEAVQPDFTVAVRELNAALRDDDDVLDLAMLTTADGVTL
LRKR"
gene
complement(8802..9884)
/genes="fkbH"
CDS
complement(8802..9884)
/genes="fkbH"
/codon_start=1
/transl_table=11
/product="fkbH"
/protein_id="AAF86387.1"
/db_xref="GI:9280389"
/translation="MTIVKCLVWDLNTLWRGTVLBDDRVVLTDREIVITLDRGI
LQVASKNDHDLAWEKLERLGVAEYFVLARI GWKPSQSVREIATELNPAFTIAPID
DQAEARVAFHLPVRCYPAPQATLLSPFSPVTSVDSRRRLMYAQFARDQA
REAYSGPDDFLRSLDLSMTIAPAGEELSVEELTSTSONMATGVHVSADLACRF
TPAHEVLVVTWDRGPHGAVGIIILEKKPSTWHLKIATSCRUVSFGAGATILNL
TQGRAGAHVLVADFRTRDRNMEIAYRFAGFADSDCFVSEVAGASAGVERLHLE
PSARPFTTLITAAIDIAPIVTSAG"
gene
complement(9895..10995)
/genes="fkbI"
CDS
complement(9895..10995)
/genes="fkbI"
/notes="acyl CoA dehydrogenase"
/codon_start=1
/transl_table=11
/product="fkbI"
/protein_id="AAF86388.1"
/db_xref="GI:9280390"
/translation="MFERDALLTDLVDRAAEWDTSGELPRDLLVRLGADGLLCAEVA
AHGGLGSLRNGEFTAHVGLSLSRVMTSQGNAWTVQRLGDAGORATFLKELT
SGKLAAVGFSEQAGSDLSAMRTRVRLDGTAVDGHKVTWTTAAAYADHLVYFGLQED
GSGAVVVPADTPGVVERVPKPSGCRAAGHADLHLDQVRVPAGAVLAGSGASLPMLV

```

```

AASLAYGRKSVAMGCVGILRCRTAAVAHARTREQFRPLGDHQLVAGHIALDLWTAEQ
IARAYCKEYASHDHWDSGSEPMVPATILAKHVAARAAGATAAQVLASAGAREGHVVE
RAYRDAKLEIMEIIEGSESCRMVLAQHALLALPA"
gene
complement(10988..11248)
/genes="fkbJ"
CDS
complement(10988..11248)
/genes="fkbJ"
/notes="ACP"
/codon_start=1
/transl_table=11
/product="fkbJ"
/protein_id="AAF86389.1"
/db_xref="GI:9280391"
/translation="MTAPVKDDPVAHALVAFLEKTKTRSDMPVDRDLFARGGLTSLFAM
ELVVLEKTFDVTIAGPDLQQLAFRTVESMVVALVHRLRAVDA"
gene
complement(11245..12093)
/genes="fkbK"
CDS
complement(11245..12093)
/genes="fkbK"
/notes="acyl CoA dehydrogenase"
/codon_start=1
Query Match 13.9%; Score 129.6; DB 1; Length 77534;
Best Local Similarity 50.7%; Pred. No. 3e-20;
Matches 374; Conservative 0; Mismatches 349; Indels 15; Gaps 2;
QY 27 CTGCGAATCGACGACCTACGCAGCTTCATTTTCAGTCGCTCAATCAGGCGACCTCACCGA 86
DB 5489 CATGAGATACGGAGTTGCACTGTTTACGGCCTTGGCCGAGCAGACAGTACGAT 5430
QY 87 AACTGCCGAAAGATTAGGCATCCCGCAGCCCACTTTTCCAGACGAATCAGCCGAGTGA 146
DB 5429 CACCGCGGAAACGGCTCAACATCTCCAGCGGACCTGTCCAGGGCCATCCGGCGGTGA 5370
QY 147 AAAACACGACGACACCCACTTTTCGACCGCGCGCGCCGCAAACTGCTCTCAACCAACG 206
DB 5369 GCGAGAGTGGCGCTGCGCTGTTTCGACCGGACCAAGAACCGGCTCCGGCTCAACAAGTA 5310
QY 207 AGGCGACGCTTCTCAACACCGCGCAGCGCATCTGTCGAGAAATTCACCTCCGCGCAAC 266
DB 5309 CGCGAGGTCTTCCGCGCCACCGCCCTGGCGCGGATCAGTAGGTCTCCGGCGCCGACGA 5250
QY 267 TGAATCAAAACGCTCATGGACCCAGAAAAAGGCAAAATCCGACTGGACTTCATGCAATC 326
DB 5249 GCGCATCACCGCCTGATCGACCCCGACACCGGCAACCGTGGCACTGGGTTTCTGCACATC 5190
QY 327 CTTGGGCACTTGGATGGTCCCGAACTTATCCGAACTTCCGCGCGGCAACACCCCAACGT 386
DB 5189 CTACGGCACCTGGCTGGTCCCGACCTCTCTCGCGGCTACCGCGGCTGGCTCCCGGCAC 5130
QY 387 AGAATTCACACTCCACCAAGCGGCGAGCAATGCTCTGCTAGATCGTGTGCTGCTGATCA 446
DB 5129 CACCTTCGAACCTGCGCGGCGACCGCCCGAGCAGCCCTGGTCGACAGCTCCGCAACAGCAG 5070
QY 447 AACTGACCTCGCATTAGTTGGCCCCCAAACTCCCGAGGTTGTCATCTTTTGGGTGGGC 506
DB 5069 GCTGGACCTGGGCTGACAGCCCAACGCCCCCGGG-----ACGATGTGAGTGGAC 5016
QY 507 GCCACTGTCTGTCAAACGACTTGGCCCTAGCTGTTTCCCGCAGATCACCGGCTGGCTTCCTT 566
DB 5015 ACCACTTGGGAGCAGGACACTGTGCTGCTCTCCACCCGCGACACCGGCTGGCC----- 4961
QY 567 TTCTGCGCAAGGAGAATTGCGCTTGNATTACTCGGCGGAGAACCTTTCGTGCGGATGCG 626
DB 4960 ----AGACGCGCGCGGTGCGGACCGCGGAGCTGCGCGATAGGCGTTCGTGCGACTGA 4905
QY 627 AGCAGGTTTGGCACCCGACTCCTCATGGATGCATTAGCCGCAAGAACGCGGTTTGTTC 686
DB 4904 GCGGCTCTCGACTGCGCCAGATCACGACCGGCTGTGCGCGCGCGCGGATTCGTGCC 4845
QY 687 CAATGTGGTTTCGAATCCATGGAACATCAACCGCTCGCAGGGCTTGTTCAGCGCGAGTCT 746
DB 4844 GCGGCTCGTCTCTGGAGAGACCGAGCTGAGCAGCTGCGGGGCCCTGGTTCGCGCGCGCT 4785

```

Qy 747 CGCGCTGGTGGTTCC 764  
||| ||| ||| |||  
Db 4784 CGGTGGCCCTCGCCC 4767

Search completed: April 13, 2005, 01:05:59  
Job time : 4175 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2005, 21:14:55 ; Search time 577 Seconds  
(without alignments)

9541.351 Million cell updates/sec

Title: US-09-826-909-1\_COPY\_232\_1161

Perfect score: 930

Sequence: 1 gggaccatgggcaacgacgg.....ggatgcattagaagagggc 930

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 930   | 100.0       | 1364   | ABA91926 | ABA91926 Coryneb   |
| 2          | 930   | 100.0       | 309400 | AAH68534 | AAH68534 C glutam  |
| 3          | 924   | 99.4        | 924    | AAH68193 | AAH68193 C glutam  |
| 4          | 903   | 97.1        | 903    | ACA01299 | ACA01299 C. Glutam |
| 5          | 439   | 47.2        | 439    | ABA91927 | ABA91927 Coryneb   |
| 6          | 131.2 | 14.1        | 77536  | AAH14651 | AAH14651 Nucleotid |
| 7          | 68    | 7.3         | 927    | ABD13720 | ABD13720 Pseudomon |
| 8          | 68    | 7.3         | 972    | ABD13492 | ABD13492 Pseudomon |
| 9          | 68    | 7.3         | 1482   | ABD13817 | ABD13817 Pseudomon |
| 10         | 61.4  | 6.6         | 987    | ABD08478 | ABD08478 Pseudomon |
| 11         | 61.4  | 6.6         | 1485   | ABD08454 | ABD08454 Pseudomon |
| 12         | 61.4  | 6.6         | 1518   | ABD08440 | ABD08440 Pseudomon |
| 13         | 60    | 6.5         | 891    | ACA43570 | ACA43570 Prokaryot |
| 14         | 60    | 6.5         | 349980 | AAH68525 | AAH68525 C glutam  |
| 15         | 59.8  | 6.4         | 954    | ACA26845 | ACA26845 Prokaryot |
| 16         | 59.2  | 6.4         | 1035   | ABD15300 | ABD15300 Pseudomon |
| 17         | 57.8  | 6.2         | 657    | ABD04404 | ABD04404 Pseudomon |
| 18         | 57.8  | 6.2         | 880    | ACA02128 | ACA02128 C. glutam |
| 19         | 57.8  | 6.2         | 882    | AAH65438 | AAH65438 C glutam  |
| 20         | 57.8  | 6.2         | 1074   | ABD04063 | ABD04063 Pseudomon |

|    |      |     |        |    |             |                    |
|----|------|-----|--------|----|-------------|--------------------|
| 21 | 57.6 | 6.2 | 1675   | 6  | ABL56944    | ABL56944 Coryneb   |
| 22 | 57.6 | 6.2 | 349980 | 5  | AAH68530    | AAH68530 C glutam  |
| 23 | 57.4 | 6.2 | 471    | 11 | ABD06737    | ABD06737 Pseudomon |
| 24 | 57.4 | 6.2 | 552    | 11 | ABD06675    | ABD06675 Pseudomon |
| 25 | 57.4 | 6.2 | 909    | 11 | ABD06700    | ABD06700 Pseudomon |
| 26 | 57.4 | 6.2 | 2592   | 11 | ABD06656    | ABD06656 Pseudomon |
| 27 | 57.2 | 6.2 | 885    | 8  | ACF74147    | ACF74147 Staphyloc |
| 28 | 57.2 | 6.2 | 31248  | 9  | ACC58250    | ACC58250 Coumermyc |
| 29 | 57.2 | 6.2 | 35359  | 9  | ACC58251    | ACC58251 Coumermyc |
| 30 | 56.6 | 6.1 | 981    | 5  | AAH67079    | AAH67079 C glutam  |
| 31 | 56.6 | 6.1 | 981    | 8  | ACA00583    | ACA00583 C. glutam |
| 32 | 56.4 | 6.1 | 771    | 11 | ABD04664    | ABD04664 Pseudomon |
| 33 | 56.4 | 6.1 | 888    | 11 | ABD04810    | ABD04810 Pseudomon |
| 34 | 56.4 | 6.1 | 1197   | 11 | ABD04732    | ABD04732 Pseudomon |
| 35 | 56   | 6.0 | 888    | 6  | ABQ69229    | ABQ69229 Listeria  |
| 36 | 56   | 6.0 | 110000 | 6  | ABQ69245_18 | Continuation (19 o |
| 37 | 56   | 6.0 | 319630 | 6  | ABQ67194_18 | Continuation (19 o |
| 38 | 54.8 | 5.9 | 711    | 11 | ABD13726    | ABD13726 Pseudomon |
| 39 | 54.8 | 5.9 | 930    | 11 | ABD13487    | ABD13487 Pseudomon |
| 40 | 54.4 | 5.8 | 1109   | 6  | ABQ68708    | ABQ68708 Listeria  |
| 41 | 54.4 | 5.8 | 3259   | 6  | ABQ70739    | ABQ70739 Listeria  |
| 42 | 54.4 | 5.8 | 110000 | 6  | ABA03041_17 | Continuation (18 o |
| 43 | 54.4 | 5.8 | 110000 | 6  | ABA03041_18 | Continuation (19 o |
| 44 | 54.2 | 5.8 | 945    | 11 | ABD03424    | ABD03424 Pseudomon |
| 45 | 54.2 | 5.8 | 984    | 11 | ABD04873    | ABD04873 Pseudomon |

#### ALIGNMENTS

#### RESULT 1

ABA91926

ID ABA91926 standard; DNA; 1364 BP.

XX ABA91926;

DT 15-MAY-2002 (first entry)

DE Corynebacterium glutamicum lysr2 gene.

KW Lysr2; transcription; regulator; L-amino acid; L-valine; L-lysine;  
KW metabolic engineering; gene; ds.

OS Corynebacterium glutamicum.

FH Key Location/Qualifiers

FT CDS 232..1164

FT /\*tag= a

FT /transl\_except= (pos:232..234, aa:Met)

FT /product= "Lysr2"

XX WO200212504-A1.

XX PD 14-FEB-2002.

XX PF 15-JUN-2001; 2001WO-EP006808.

XX PR 10-AUG-2000; 2000DE-01039047.

XX PR 03-MAR-2001; 2001DE-01010346.

XX (DEGS ) DEGUSSA AG.

XX Moeckel B, Farwick M, Hermann T, Kreutzer C, Pfefferle W;

XX WPI; 2002-227155/28.

XX P-PSDB; AAM51006.

XX Novel lysr2 gene of coryneform bacteria encoding Lysr2 protein which is a  
transcription regulator, useful for fermentative production of L-lysine  
and L-valine and as a probe detecting polynucleotides encoding Lysr2.

XX Claim 5; Page 36-38; 44pp; English.

CC The present sequence is that of the novel *lysR2* gene of *Corynebacterium*  
 CC glutamicum strain ATCC 13032. The gene codes for *lysR2* protein (see  
 CC AAM51006), a transcriptional regulator of the *lysR* family. The gene was  
 CC identified in a cosmid gene library of ATCC 13032 using algorithms and  
 CC sequence analysis programs. Vector pCR2.1<sub>lysR2int</sub>, which carries an  
 CC internal fragment (see ABA91927) of the *lysR2* gene, is claimed, and is  
 CC deposited in Escherichia coli strain TOP10F/pCR2.1<sub>lysR2int</sub> as DSM 13617.  
 CC Also claimed are coryneform bacteria in which the *lysR2* gene is  
 CC attenuated, preferably eliminated, especially by deletion. These bacteria  
 CC are used for the production of L-amino acids. Insertional mutagenesis of  
 CC the *lysR2* gene in the lysine producer *C. glutamicum* DSM 5715 and the  
 CC valine producer *Brevibacterium lactofermentum* FERM BP-1763 using vector  
 CC pCR2.1<sub>lysR2int</sub> improved production of L-lysine and L-valine,  
 CC respectively. Further genes of L-amino acid biosynthetic pathways may  
 CC also be manipulated. Polynucleotide sequences from the *lysR2* gene may be  
 CC used as hybridisation probes, e.g. in arrays, microarrays or DNA chips,  
 CC for discovering related sequences

XX  
 SQ Sequence 1364 BP; 291 A; 381 C; 370 G; 322 T; 0 U; 0 Other;

Query Match 100.0%; Score 930; DB 6; Length 1364;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-264;  
 Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACCATGGGCAACGACGGCGGAGACCTGCGAATCGACGACTACGACGCTTCAATTCA 60  
 DB 232 GTGACCATGGGCAACGACGGCGGAGACCTGCGAATCGACGACTACGACGCTTCAATTCA 291  
 QY 61 GTGCGTCAATCAGGCACCTCAACGAACTGCGGAAAGATTAGGCATCCGCGAGCCACA 120  
 DB 292 GTGCGTCAATCAGGCACCTCAACGAACTGCGGAAAGATTAGGCATCCGCGAGCCACA 351  
 QY 121 CTTTCCAGACGATCAGCGAGTGGAAACACGAGGACCCCACTTTTCGACGCGCC 180  
 DB 352 CTTTCCAGACGATCAGCGAGTGGAAACACGAGGACCCCACTTTTCGACGCGCC 411  
 QY 181 GGCCGCAAACTCGTCTCAACCAACGAGGCGACGCTTCTCAACGCGCAGCGCCATC 240  
 DB 412 GGCCGCAAACTCGTCTCAACCAACGAGGCGACGCTTCTCAACGCGCAGCGCCATC 471  
 QY 241 GTGCGAGAAATCAATCCGCGCCCAACTGAAATCAACGCTCATCGACCCAGAAAAGGC 300  
 DB 472 GTGCGAGAAATCAATCCGCGCCCAACTGAAATCAACGCTCATCGACCCAGAAAAGGC 531  
 QY 301 ACAATCCGACTGGATTCATGATTCCTTGGGCACTTGGATGGTCCCGACTTATCGA 360  
 DB 532 ACAATCCGACTGGATTCATGATTCCTTGGGCACTTGGATGGTCCCGACTTATCGA 591  
 QY 361 ACATTCCGCGCGCAACACCCCAACCTAGAAATTCCAACTCCACAGCGCGCAGCAATGCTC 420  
 DB 592 ACATTCCGCGCGCAACACCCCAACCTAGAAATTCCAACTCCACAGCGCGCAGCAATGCTC 651  
 QY 421 CTGGTAGATCGTGTGTTTGGCTGATGAAACTGACCTCGCATTTAGTTGGCCCAAACTGCC 480  
 DB 652 CTGGTAGATCGTGTGTTTGGCTGATGAAACTGACCTCGCATTTAGTTGGCCCAAACTGCC 711  
 QY 481 GAGTTGTGACTCTTTAGGTTGGGCGCCACTGCTTCGTCAACGACTTGCCTACTGTT 540  
 DB 712 GAGTTGTGACTCTTTAGGTTGGGCGCCACTGCTTCGTCAACGACTTGCCTACTGTT 771  
 QY 541 CCCGAGATCACCGGCTTGCCTCTTTTTCGCCAAGGAGAAATTCGCTGTGATTAATCGG 600  
 DB 772 CCCGAGATCACCGGCTTGCCTCTTTTTCGCCAAGGAGAAATTCGCTGTGATTAATCGG 831  
 QY 601 GCGGAAGAACCTTTTCGTGCGCATCGAGAGGTTTCGGACCCGACTCTCATGATGCA 660  
 DB 832 GCGGAAGAACCTTTTCGTGCGCATCGAGAGGTTTCGGACCCGACTCTCATGATGCA 891  
 QY 661 TTAGCCGAGAGACCGGTTTTCCTCCATGTTGTTTTCGATTCATCGACTCCACACC 720  
 DB 892 TTAGCCGAGAGACCGGTTTTCCTCCATGTTGTTTTCGATTCATCGACTCCACACC 951  
 QY 721 GTCCGAGGCGTTGTGACGCGCAGGTCTCGGCGTTGGTGTGTTTCCGATGATCCGTAC 780

DB 952 GTCCAGGCTGTTCAGCGCAGGCTCTGGCGTGTGGTGGTTCGATGATGATCCGTAC 1011  
 QY 781 CTTCCACAGTGGGAATCGTGCACGCCCACTTAGTCCACCGCTTATAGGGAACCTAGGT 840  
 DB 1012 CTTCCACAGTGGGAATCGTGCACGCCCACTTAGTCCACCGCTTATAGGGAACCTAGGT 1071  
 QY 841 TTGGTGTGGCAGTCAACGCGGCGCGGACCTCGGTGGATAACTTCCGGAAGTTCGTG 900  
 DB 1072 TTGGTGTGGCAGTCAACGCGGCGCGGACCTCGGTGGATAACTTCCGGAAGTTCGTG 1131  
 QY 901 CGCGGATCGAGGTATGCATTAGAGAGGGC 930  
 DB 1132 CGCGGATCGAGGTATGCATTAGAGAGGGC 1161

RESULT 2  
 AAH68534/c  
 ID AAH68534 standard; DNA; 309400 BP.  
 XX  
 AC AAH68534;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE C glutamicum coding sequence fragment SEQ ID NO: 7069.  
 XX  
 KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
 XX  
 OS organic acid synthesis; ds.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN EP1108790-A2.  
 XX  
 PD 20-JUN-2001.  
 XX  
 PF 18-DEC-2000; 2000EP-00127688.  
 XX  
 PR 16-DEC-1999; 99JP-00377484.  
 PR 07-APR-2000; 2000JP-00159162.  
 PR 03-AUG-2000; 2000JP-00280988.  
 XX  
 PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX  
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 XX  
 DR WPI; 2001-376931/40.  
 XX  
 PS Novel polynucleotides derived from Coryneform bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analyzing  
 PT expression profile or pattern of a gene and identifying homologous gene.  
 XX  
 PS Disclosure; SEQ ID NO 7069; 246pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of nucleotide and protein  
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of coryneform bacterium, measuring expression amount and analyzing  
 CC the expression profile or expression pattern of a gene derived from  
 CC Coryneform bacterium, and identifying a homologue of a gene derived from  
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino  
 CC acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a nucleic acid described  
 CC in the exemplification of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from the European Patent Office

SQ Sequence 309400 BP; 70133 A; 86477 C; 83115 G; 69675 T; 0 U; 0 Other;

Query Match 100.0%; Score 930; DB 5; Length 309400;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-263;  
 Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
QY 1 GTGACCATGGGCAACGACGGGGAGACCTTGGGAATCGACGACCTTACAGCTTCAATTTCA 60
Db 128423 GTGACCATGGGCAACGACGGGGAGACCTTGGGAATCGACGACCTTACAGCTTCAATTTCA 128364
QY 61 GTCGCTCAATCAGGCGACCTCACCGAACTGCGGAAAGATTAGGCATCCGCGAGCCACA 120
Db 128363 GTCGCTCAATCAGGCGACCTCACCGAACTGCGGAAAGATTAGGCATCCGCGAGCCACA 128304
QY 121 CTTTCCAGACGAATCAGCGAGTGGAAACACACGCGAGGACCCCACTTTTCGACCGCGCC 180
Db 128303 CTTTCCAGACGAATCAGCGAGTGGAAACACACGCGAGGACCCCACTTTTCGACCGCGCC 128244
QY 181 GCGCCGAACTCGTCTCTCAACCAACGAGCGCACGCTTCTTCTCAACACGCGAGCGCCATC 240
Db 128243 GCGCCGAACTCGTCTCTCAACCAACGAGCGCACGCTTCTTCTCAACACGCGAGCGCCATC 128184
QY 241 GTCGCGAGAAATCAACTCGCGCGCACTGAAATCAACCGCTCATGAGCCCGAGAAAAGGC 300
Db 128183 GTCGCGAGAAATCAACTCGCGCGCACTGAAATCAACCGCTCATGAGCCCGAGAAAAGGC 128124
QY 301 ACAATCCGACTGCACTTCATGCATTCCTTGGGCACTTGGATGGTCCCGAACTTATCCGA 360
Db 128123 ACAATCCGACTGCACTTCATGCATTCCTTGGGCACTTGGATGGTCCCGAACTTATCCGA 128064
QY 361 ACAATCCGCGCGCAACACCCCAACGTAGAAATCCAACTCCACGAGCGCAGCAATGCTC 420
Db 128063 ACAATCCGCGCGCAACACCCCAACGTAGAAATCCAACTCCACGAGCGCAGCAATGCTC 128004
QY 421 CTGGTAGATCGTGTGTTTGGCTGATGAATGACCTCGCATTTAGTTGGGCCCAAACTGCG 480
Db 128003 CTGGTAGATCGTGTGTTTGGCTGATGAATGACCTCGCATTTAGTTGGGCCCAAACTGCG 127944
QY 481 GAGGTTGGTACCTCTTAGGTGGCGCCGCACTGCTTGGTCAACGACTTGGCCCTAGCTGT 540
Db 127943 GAGGTTGGTACCTCTTAGGTGGCGCCGCACTGCTTGGTCAACGACTTGGCCCTAGCTGT 127884
QY 541 CCGCGAGATCACCGGCTTGCCTCTTTCTGCGCAAGGAGAAATGCCGTTGATTACTGCG 600
Db 127883 CCGCGAGATCACCGGCTTGCCTCTTTCTGCGCAAGGAGAAATGCCGTTGATTACTGCG 127824
QY 601 GCGGAGAACCTTTGTGGCGATGCGAGCAGTTTCGGACCGGACTCTCTCATGTATGCA 660
Db 127823 GCGGAGAACCTTTGTGGCGATGCGAGCAGTTTCGGACCGGACTCTCTCATGTATGCA 127764
QY 661 TTAGCGGAGAACCGGTTTGTTCCTCAATGTGTTTTCGAATCCATGGAATCCACCAC 720
Db 127763 TTAGCGGAGAACCGGTTTGTTCCTCAATGTGTTTTCGAATCCATGGAATCCACCAC 127704
QY 721 GTCGCGAGGCTTGTTCAGCGCAGGTCTCGGCGTTGTTGTTCCGATGGATGATCCGTAC 780
Db 127703 GTCGCGAGGCTTGTTCAGCGCAGGTCTCGGCGTTGTTGTTCCGATGGATGATCCGTAC 127644
QY 781 CTTTCCACAGTGGGAATCGTCAACCGCCACTTAGTTCACCGCTTATAGGAACTAGGT 840
Db 127643 CTTTCCACAGTGGGAATCGTCAACCGCCACTTAGTTCACCGCTTATAGGAACTAGGT 127584
QY 841 TTGGTGTGCGCACTCAACCGGGCGCGCACTCGGTGGATAACTTCGCGAAGTTCTGTG 900
Db 127583 TTGGTGTGCGCACTCAACCGGGCGCGCACTCGGTGGATAACTTCGCGAAGTTCTGTG 127524
QY 901 GCGGGATCGAGGTATGCAATTAGAAAGAGGCG 930
Db 127523 GCGGGATCGAGGTATGCAATTAGAAAGAGGCG 127494
```

## RESULT 3

AAH68193

ID AAH68193 standard; DNA; 924 BP.

XX

AC AAH68193;

XX

DT 26-SEP-2001 (first entry)

XX

```
DE C glutamicum coding sequence fragment SEQ ID NO: 3228.
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
XX Corynebacterium glutamicum.
XX EP1108790-A2.
XX 20-JUN-2001.
XX 18-DEC-2000; 2000EP-00127688.
XX 16-DEC-1999; 99JP-00377484.
XX 07-APR-2000; 2000JP-00159162.
XX 03-AUG-2000; 2000JP-00280988.
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
XX P-PSDB; AAG92974.
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX Claim 8; SEQ ID NO 3228; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein
XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of coryneform bacterium, measuring expression amount and analysing
XX the expression profile or expression pattern of a gene derived from
XX Coryneform bacterium, and identifying a homologue of a gene derived from
XX coryneform bacterium. Coryneform bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a nucleic acid described
XX in the exemplification of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the European Patent Office
XX Sequence 924 BP; 207 A; 286 C; 239 G; 192 T; 0 U; 0 Other;
XX Query Match 99.4%; Score 924; DB 5; Length 924;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-262;
XX Matches 924; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 ATGGGCAACGACGGCGGAGACCTGCGAATCGACGACCTACGAGCTTCATTTTCAGTCGCT 66
Db 1 ATGGGCAACGACGGCGGAGACCTGCGAATCGACGACCTACGAGCTTCATTTTCAGTCGCT 60
QY 67 CAATCAGGCCACCTCACCGAAACTCCGAAAGATTAGGCATCCGCGAGCCCACTTTCC 126
Db 61 CAATCAGGCCACCTCACCGAAACTCCGAAAGATTAGGCATCCGCGAGCCCACTTTCC 120
QY 127 AGACGAATCAGCGAGTGGAAAAACACGCGAGCAGCCCACTTTTCGACCGCGCGCGCG 186
Db 121 AGACGAATCAGCGAGTGGAAAAACACGCGAGCAGCCCACTTTTCGACCGCGCGCGCG 180
QY 187 AAATCTCGTCTCAACCAACGAGGCGCACGCTTCTCAACGCGCAGCGCCCATCGTCGCA 246
Db 181 AAATCTCGTCTCAACCAACGAGGCGCACGCTTCTCAACGCGCAGCGCCCATCGTCGCA 240
QY 247 GAATTCACCTCCGCGCAACTGAAATCAAAACGCTCATGAGACCCAGAAAAGCAATC 306
Db 241 GAATTCACCTCCGCGCAACTGAAATCAAAACGCTCATGAGACCCAGAAAAGCAATC 300
QY 307 CGACTGGACTTCATGCATTCCTTGGGCACTTGGTGGTCCCGCACTTATCCGAACATTC 366
Db 301 CGACTGGACTTCATGCATTCCTTGGGCACTTGGTGGTCCCGCACTTATCCGAACATTC 360
```

QY 367 CGCGCGAACACCCAAACGTAGAAATTCACCTCCACCAAGCGGACGCAATGCTCTGTGTA 426  
DB 361 CGCGCGAACACCCAAACGTAGAAATTCACCTCCACCAAGCGGACGCAATGCTCTGTGTA 420  
QY 427 GATCGTGTGTTGGCTGATGAACACTGACCTCGCAATGTTGGCCCAAACTCCCGAGGTT 486  
DB 421 GATCGTGTGTTGGCTGATGAACACTGACCTCGCAATGTTGGCCCAAACTCCCGAGGTT 480  
QY 487 GGTACTCTTTAGGTGGCGGCACCTGCTTCGTCGTAACGACTTGCCTAGCTGTTCCCGCA 546  
DB 481 GGTACTCTTTAGGTGGCGGCACCTGCTTCGTCGTAACGACTTGCCTAGCTGTTCCCGCA 540  
QY 547 GATCACCGGCTTGCCTCTCTTCTCGCCCAAGGAGAAATTCGCTTACTGCGGCGGAA 606  
DB 541 GATCACCGGCTTGCCTCTCTTCTCGCCCAAGGAGAAATTCGCTTACTGCGGCGGAA 600  
QY 607 GAACCTTTTGTGCGGATCGAGCAGGTTTCGGCACCCGACTCTCATGATGCAATAGCC 666  
DB 601 GAACCTTTTGTGCGGATCGAGCAGGTTTCGGCACCCGACTCTCATGATGCAATAGCC 660  
QY 667 GAAGAAGCGGTTTGTTCGCAATGTGTTTTCGAAATCGATGGAACCTCACACCGTCGGA 726  
DB 661 GAAGAAGCGGTTTGTTCGCAATGTGTTTTCGAAATCGATGGAACCTCACACCGTCGGA 720  
QY 727 GGGCTTGTACGCGAGGTTCTGGCGTGTGTTGTTCCGATGGATGATCCGTTACCTTCCC 786  
DB 721 GGGCTTGTACGCGAGGTTCTGGCGTGTGTTGTTCCGATGGATGATCCGTTACCTTCCC 780  
QY 787 ACAGTGGGAATCGTCAACGCCCACTTAGTCCACCCGCTTATAGGGAACCTAGGTTTGGTG 846  
DB 781 ACAGTGGGAATCGTCAACGCCCACTTAGTCCACCCGCTTATAGGGAACCTAGGTTTGGTG 840  
QY 847 TGGCACTCAACGCGGGCGGCACCTCGGTFGATATACTTCGGAAGTTTGTGCGGGA 906  
DB 841 TGGCACTCAACGCGGGCGGCACCTCGGTFGATATACTTCGGAAGTTTGTGCGGGA 900  
QY 907 TCGAGGTATGCAATAGAGAGGCG 930  
DB 901 TCGAGGTATGCAATAGAGAGGCG 924

## RESULT 4

ACA01299

ID ACA01299 standard; DNA; 903 BP.

XX ACA01299;

AC ACA01299;

XX 03-JUN-2003 (first entry)

XX C. glutamicum derived ORF SEQ ID 1290.

XX Coryneform; nucleic acid array; fermentation; culture; ds.

XX Corynebacterium glutamicum.

XX DE10128510-A1.

XX 19-DEC-2002.

XX 13-JUN-2001; 2001DE-01028510.

XX 13-JUN-2001; 2001DE-01028510.

XX (DEGS ) DEGUSSA AG.

XX Farwick M, Moeckel B, Pfeifferle W, Bathe B, Huthmacher K;

XX WPI; 2003-279970/28.

XX New nucleic acid array useful for monitoring mRNA expression of  
XX Corynebacterium glutamicum during fermentation, comprising nucleic acid  
XX from Corynebacterium glutamicum.

XX

PS Claim 1; Page 457-458; 709pp; German.

XX This invention describes a novel nucleic acid array involving  
CC Corynebacterium glutamicum polynucleotides. The arrays are used to  
CC analyse C. glutamicum, particularly for monitoring a fermentation process  
CC to determine expression levels of C. glutamicum cellular mRNA. Such  
CC monitoring particularly differentiates between expression levels of  
CC different strains of C. glutamicum and allows the adjustment of different  
CC culture and fermentation conditions. ACA0010-ACA02188 represent C.  
CC glutamicum derived polynucleotides described in the disclosure of the  
XX invention

SQ Sequence 903 BP; 201 A; 281 C; 230 G; 191 T; 0 U; 0 Other;  
Query Match 97.1%; Score 903; DB 8; Length 903;  
Best Local Similarity 100.0%; Pred. No. 2.7e-256;  
Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 CTGCGAATCGAGCACTACGCGCTTCATTTTCAGTCGCTCAATCAGGCGCACTCACCGAA 87  
DB 1 CTGCGAATCGAGCACTACGCGCTTCATTTTCAGTCGCTCAATCAGGCGCACTCACCGAA 60  
QY 88 ACTGCGCAAGATTTAGGCATCCCGCAGCCCACTTTTCCAGACCAATCAGCGGAGTGAA 147  
DB 61 ACTGCGCAAGATTTAGGCATCCCGCAGCCCACTTTTCCAGACCAATCAGCGGAGTGAA 120  
QY 148 AAAACGCGAGCACTTTCGCGCGCGCGCGCAAACTCGTCTCAACCAACGA 207  
DB 121 AAAACGCGAGCACTTTCGCGCGCGCGCGCAAACTCGTCTCAACCAACGA 180  
QY 208 GGGCAGGCTTCTCAACCGCAGCGCCATCTCGCAGNAATTCAACTCCGCGCACT 267  
DB 181 GGGCAGGCTTCTCAACCGCAGCGCCATCTCGCAGNAATTCAACTCCGCGCACT 240  
QY 268 GAAATCAAAAGCTTCATGGACCCAGAAAAGGCAAACTCCGACCTGGAATTCATGCAATCC 327  
DB 241 GAAATCAAAAGCTTCATGGACCCAGAAAAGGCAAACTCCGACCTGGAATTCATGCAATCC 300  
QY 328 TTGGGCACTTGGATGGTCCCGCAACTTATCCGAACTTCGCGCGCGCAACCCCAACGTA 387  
DB 301 TTGGGCACTTGGATGGTCCCGCAACTTATCCGAACTTCGCGCGCGCAACCCCAACGTA 360  
QY 388 GAATTCCAACTCCACCAAGCGGCAATGCTCTCGTGTAGATCGTGTGTTGGCTGATGAA 447  
DB 361 GAATTCCAACTCCACCAAGCGGCAATGCTCTCGTGTAGATCGTGTGTTGGCTGATGAA 420  
QY 448 ACTGACCTCGCATTTAGTTGGCCCAAACTCCGCGAGGTTGGTACTCTTTAGGGTGGCG 507  
DB 421 ACTGACCTCGCATTTAGTTGGCCCAAACTCCGCGAGGTTGGTACTCTTTAGGGTGGCG 480  
QY 508 CCACTGCTTGTCAACGACTTGCCTAGCTGTTCCCGCAGATCAACCGGCTTGCCTCTTT 567  
DB 481 CCACTGCTTGTCAACGACTTGCCTAGCTGTTCCCGCAGATCAACCGGCTTGCCTCTTT 540  
QY 568 TCTGGCCAAAGAGAAATTTGCGGTTGATTACTCGCGCGGGAAGAACCTTTCGTGGCGATGGA 627  
DB 541 TCTGGCCAAAGAGAAATTTGCGGTTGATTACTCGCGCGGGAAGAACCTTTCGTGGCGATGGA 600  
QY 628 GCAGGTTTTCGGCACCCGACTCTCTCATGGATGCAATAGCCGAAGAGCGGTTTGTGTTCCC 687  
DB 601 GCAGGTTTTCGGCACCCGACTCTCTCATGGATGCAATAGCCGAAGAGCGGTTTGTGTTCCC 660  
QY 688 AATGTGTTTTCGAATCCATGGAACTCACCACCGTCGAGGCTTGTTCAGGCGAGGTC 747  
DB 661 AATGTGTTTTCGAATCCATGGAACTCACCACCGTCGAGGCTTGTTCAGGCGAGGTC 720  
QY 748 GGGCTTGTGTTTCCGATGGATGATCCGTTACCTTCCACAGTGGGAATCGTCAACGC 807  
DB 721 GGGCTTGTGTTTCCGATGGATGATCCGTTACCTTCCACAGTGGGAATCGTCAACGC 780  
QY 808 CCACTTAGTCCACCGCTTATAGGGAATAGGTTTGGTGGCGCACTCAACCGCGGGCG 867

```

Db 781 CCACCTTAGTCCACCCCGCTTATAGGGAACCTAGGTTGGTGTGCGACTCAACCGGGCG 840
QY 868 GCACCTGCGGTGATAAATTCGGAAGTTTCGGGGGATCGAGGTATGATAGAGAG 927
Db 841 GCACCTGCGGTGATAAATTCGGAAGTTTCGGGGGATCGAGGTATGATAGAGAG 900
QY 928 GGC 930
Db 901 GGC 903

RESULT 5
ABA91927
ID ABA91927 standard; DNA; 439 BP.
XX
AC ABA91927;
DT 15-MAY-2002 (first entry)
XX
DE Corynebacterium glutamicum lysr2 gene internal fragment.
XX
KW Lysr2; transcription; regulator; L-amino acid; L-valine; L-lysine;
KW metabolic engineering; pCR2.1lysR2int; gene; ds.
XX
OS Corynebacterium glutamicum.
XX
PN WO200212504-A1.
XX
PD 14-FEB-2002.
XX
PF 15-JUN-2001; 2001WO-BF006808.
XX
PR 10-AUG-2000; 2000DE-01039047.
XX
PR 03-MAR-2001; 2001DE-01010346.
XX
PA (DEGS ) DEGUSSA AG.
XX
PI Moeckel B, Farwick M, Hermann T, Kreutzer C, Pfefferle W;
XX
WPI; 2002-227155/28.
XX
DR Novel lysr2 gene of coryneform bacteria encoding Lysr2 protein which is a
PT transcription regulator, useful for fermentative production of L-lysine
PT and L-valine and as a probe detecting polynucleotides encoding Lysr2.
XX
PS Example 3; Page 39; 44pp; English.
XX
CC The present sequence is that of an internal fragment of the novel lysr2
CC gene of Corynebacterium glutamicum strain ATCC 13032. The gene fragment
CC was obtained by PCR amplification of chromosomal DNA using the primers
CC given in ABA91928 and ABA91929. It was used in the construction of
CC claimed vector pCR2.1lysR2int, which is deposited in Escherichia coli
CC TOP10F as DSM 13617. The full-length lysr2 gene (see ABA91926) codes for
CC Lysr2 protein (see AM51006), a transcription regulator of the lysr
CC family. The invention provides coryneform bacteria in which the lysr2
CC gene is attenuated, preferably eliminated, especially by deletion. These
CC bacteria are used in the production of L-amino acids. In an example of
CC this, insertional mutagenesis of the lysr2 gene in the lysine producer C.
CC glutamicum DSM 5715 and in the valine producer Brevibacterium
CC lactofermentum FERM BP-1763 using pCR2.1lysR2int improved yields of L-
CC lysine and L-valine, respectively
XX
SQ Sequence 439 BP; 100 A; 135 C; 106 G; 98 T; 0 U; 0 Other;

Query Watch 47.2%; Score 439; DB 6; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.3e-119;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 CCATCGTCCGAGAAATTCACCTCGCGCGCAACTGAAATCAAAAGCCCTCATGGACCCAGAAA 295
Db 1 CCATCGTCCGAGAAATTCACCTCGCGCGCAACTGAAATCAAAAGCCCTCATGGACCCAGAAA 60
QY 296 AAGGCACAATCCGACTGGGACTTTCATGCAATTCCTTGGGCACTTGGATGGTCCCGAACTT 355

```

```

Db 61 AAGGCACAATCCGACTGGGACTTTCATGCAATTCCTTGGGCACTTGGATGGTCCCGAACTT 120
QY 356 TCCGAACATTCGCGCGCGCAACACCAACGCTAGAATTCCAACTCCACCAAGCGGAGCAA 415
Db 121 TCCGAACATTCGCGCGCGCAACACCAACGCTAGAATTCCAACTCCACCAAGCGGAGCAA 180
QY 416 TGCTCCTGCTAGATCGTGTGTTTGGCTGATGAACCTGACCTCGCATTAGTTGGCCCCAAC 475
Db 181 TGCTCCTGCTAGATCGTGTGTTTGGCTGATGAACCTGACCTCGCATTAGTTGGCCCCAAC 240
QY 476 CTGCGCGAGTTGGTACCTCTTTAGGTTGGGCGCCACTGCTTCAACGACTTGCCTAG 535
Db 241 CTGCGCGAGTTGGTACCTCTTTAGGTTGGGCGCCACTGCTTCAACGACTTGCCTAG 300
QY 536 CTGTTCCCGCAGATCACCGGCTTGCTCTCTTTCTGGCAAGAGAAATTCGCGTTGATTA 595
Db 301 CTGTTCCCGCAGATCACCGGCTTGCTCTCTTTCTGGCAAGAGAAATTCGCGTTGATTA 360
QY 596 CTGCGCGCGAAGAACCTTTTCGTGGCGATGCGAGCAGGTTTCGGCACCCGACTCCTCATGG 655
Db 361 CTGCGCGCGAAGAACCTTTTCGTGGCGATGCGAGCAGGTTTCGGCACCCGACTCCTCATGG 420
QY 656 ATGCATTAGCCGAGAAGC 674
Db 421 ATGCATTAGCCGAGAAGC 439

RESULT 6
AAA14651/c
ID AAA14651 standard; DNA; 77536 BP.
XX
AC AAA14651;
DT 08-AUG-2000 (first entry)
XX
DE Nucleotide sequence of the FK-520 biosynthetic gene cluster.
XX
KW FK-520; polyketide synthase; PKS; gene cluster; immunosuppressant;
KW Streptomyces hygroscopicus var. ascomyceticus; immunophilin;
KW FK-506 binding protein; polyketide compound; transplant rejection;
KW graft-versus-host disease; uveitis; alopecia universalis;
KW autoimmune chronic active hepatitis; inflammatory bowel disease;
KW multiple sclerosis; primary biliary cirrhosis; scleroderma;
KW neurite outgrowth; nerve regrowth; Parkinson's disease;
KW Alzheimer's disease; stroke; traumatic spinal cord; brain injury;
KW peripheral neuropathy; ss.
XX
OS Streptomyces hygroscopicus.
XX
FH Key Location/Qualifiers
CDS complement(412..1836)
FT /tag= a
FT /note= "fkBW gene"
FT complement(2020..3579)
FT /tag= b
FT /note= "fkBV gene"
FT 3969..4496
FT /tag= c
FT /note= "fkBR2 gene"
FT complement(4595..5488)
FT /tag= d
FT /note= "fkBR1 gene"
FT 5601..6818
FT /tag= e
FT /note= "fkBE gene"
FT 6808..8052
FT /tag= f
FT /note= "fkBF gene"
FT 8156..8824
FT /tag= g
FT /note= "fkBG gene"
FT complement(9122..9883)

```

```

FT      CDS      /*tag= h
FT      /note= "fkfH gene"
FT      complement(9894..10994)
FT      /*tag= i
FT      /note= "fkfI gene"
FT      complement(10987..11247)
FT      /*tag= j
FT      /note= "fkfJ gene"
FT      complement(11244..12092)
FT      /*tag= k
FT      /note= "fkfK gene"
FT      complement(12113..13150)
FT      /*tag= l
FT      /note= "fkfL gene"
FT      complement(13212..23988)
FT      /*tag= m
FT      /note= "fkfC gene"
FT      complement(13452..13662)
FT      /*tag= n
FT      /note= "ACPE"
FT      complement(13761..14394)
FT      /*tag= o
FT      /note= "KR6"
FT      complement(14517..15294)
FT      /*tag= p
FT      /note= "ER6"
FT      complement(15438..16587)
FT      /*tag= q
FT      /note= "dehydratase domain (DH) 6"
FT      complement(16587..17820)
FT      /*tag= r
FT      /note= "acyltransferase domain (AT) 6"
FT      complement(17820..19053)
FT      /*tag= s
FT      /note= "KS6"
FT      complement(19116..19326)
FT      /*tag= t
FT      /note= "ACPS"
FT      complement(19464..20097)
FT      /*tag= u
FT      /note= "KRS"
FT      complement(20241..21420)
FT      /*tag= v
FT      /note= "DH5"
FT      complement(21420..22653)
FT      /*tag= w
FT      /note= "AT5"
FT      complement(22653..23892)
FT      /*tag= x
FT      /note= "KS5"
FT      complement(23992..46573)
FT      /*tag= y
FT      /note= "fkfB gene"
FT      complement(24163..24373)
FT      /*tag= z
FT      /note= "ACP4"
FT      complement(24997..26146)
FT      /*tag= aa
FT      /note= "DH4 (inactive)"
FT      complement(26146..27430)
FT      /*tag= ab
FT      /note= "AT4"
FT      complement(27430..28684)
FT      /*tag= ac
FT      /note= "KS4"
FT      complement(28750..28960)
FT      /*tag= ad
FT      /note= "ACP3"
FT      complement(28750..28960)
FT      /*tag= al
FT      /note= "DH2 (inactive)"
FT      complement(29092..29740)
FT      /*tag= ae

FT      misc_feature  /note= "KR3"
FT      complement(29869..31018)
FT      /*tag= af
FT      /note= "DH3 (inactive)"
FT      complement(31018..32185)
FT      /*tag= ag
FT      /note= "AT3"
FT      complement(32185..33439)
FT      /*tag= ah
FT      /note= "KR3"
FT      complement(33505..33715)
FT      /*tag= ai
FT      /note= "ACP2"
FT      complement(33823..34480)
FT      /*tag= aj
FT      /note= "KR2"
FT      complement(34606..35749)
FT      /*tag= ak
FT      complement(35749..37144)
FT      /*tag= am
FT      /note= "AT2"
FT      complement(37145..38296)
FT      /*tag= an
FT      /note= "KS2"
FT      complement(38371..38581)
FT      /*tag= ao
FT      /note= "ACP1"
FT      complement(38677..39307)
FT      /*tag= ap
FT      /note= "KR1"
FT      complement(39442..40609)
FT      /*tag= aq
FT      /note= "DH1"
FT      complement(40609..41842)
FT      /*tag= ar
FT      /note= "AT1"
FT      complement(41842..43093)
FT      /*tag= as
FT      /note= "KS of extender module 1 (KS1)"
FT      complement(43144..43660)
FT      /*tag= at
FT      /note= "ACP of loading domain"
FT      complement(43777..44629)
FT      /*tag= au
FT      /note= "ER of loading domain"
FT      complement(44974..46573)
FT      /*tag= av
FT      /note= "CoA ligase of loading domain"
FT      46754..47788
FT      /*tag= aw
FT      /note= "fkfO gene"
FT      47785..52272
FT      /*tag= ax
FT      /note= "fkfP gene"
FT      52275..71465
FT      /*tag= ay
FT      /note= "fkfA gene"
FT      52362..53576
FT      /*tag= az
FT      /note= "KR7"
FT      53577..54716
FT      /*tag= bb
FT      /note= "AT7"
FT      54717..55871
FT      /*tag= ba
FT      /note= "DH7"
FT      56019..56819
FT      /*tag= bc
FT      /note= "ER7"
FT      56943..57575
FT      /*tag= bd
FT      /note= "KR7"
FT      57710..57920

```

```

FT      /*tag= be
FT      /note= "ACP7"
FT      misc_feature 5790..59243
FT      /*tag= bf
FT      /note= "KS8"
FT      misc_feature 59244..60398
FT      /*tag= bg
FT      /note= "AT8"
FT      misc_feature 60399..61412
FT      /*tag= bh
FT      /note= "DH8 (inactive)"

Query Match      14.1%; Score 131.2; DB 3; Length 77536;
Best Local Similarity 50.8%; Pred. No. 2.9e-27;
Matches 375; Conservative 0; Mismatches 348; Indels 15; Gaps 2;

QY      27  CTTGCGAATCGACGACCTTACGAGCTTCAATTCAGTCGCTCAATCAGGCCACTCACCGA 86
Db      |||
Db      5489 CATGGAGATACGGGAGTTCAGTGGTTACGGCTTGGCGGACGACGACGACGATGACGAT 5430
QY      87  AACTGCCGAAAGATTAGGCATCCGCGAGCCACACATTTCCAGACGAATCAGCCGAGTGGA 146
Db      |||
Db      5429 CACCGCGGAACGGCTCAACATCTCCAGCGACCTGTCCAGGGCCATCCGGGGCTGA 5370
QY      147 AAAACAGCAGGACACCCCACTTTTCGACCGCGCGCGCAAACTGCTCTCAACCAACG 206
Db      |||
Db      5369 GCGACGAGTGGCGCTGCTGTTGACCGGGACAGAACCGGCTCCGGCTCAACAAGTA 5310
QY      207 AGCCACGCTTCTCAACAGCGCAGCGCATCTGCGCAGAAATCAACTCGCGCGCAAC 266
Db      |||
Db      5309 CGCGAGAGTCTTCCGCGCCACGCGCTGCGCGCATCAGTGAAGTCTCCGGCGCGAGCA 5250
QY      267 TGAATCAAAAGCGCTCATGGACCCAGAAAAGCAATCCGACTGCTTCATGCTATTC 326
Db      |||
Db      5249 GCGCATACGCGCTGATGACCCCGACACCGGACCGTGGCACTGGTTCTGCACTC 5190
QY      327 CTTGGCACTTGAATGGTCCCGAACTTATTCGAACTTCCGCGCGCAACACCCCAACGT 386
Db      |||
Db      5189 CTACGSCACTGGCTGGTGGCGGACCTCTCGCGGGTACCGCGGCTCGGCTCCCGGCAC 5130
QY      387 AGAATTCGAATCCAGCGGCGGACGATGCTCTGCTGATGATCGTGTTCGCTGATGA 446
Db      |||
Db      5129 CACCTTCGAATCGCGGCGGACGCGCGCGGCTGCTGCGACGCTCCGCAACAGCAG 5070
QY      447 AACTGACCTCGCATTTAGTTGGCCCAACCTGCGGAGTTGTTACCTCTTTAGGGTGGSC 506
Db      |||
Db      5069 GCTGGACCTGGGCTGACAGCCCAACGACCGCGCGGG-----ACGATGTGAGTGGAC 5016
QY      507 GCCACTGCTTCTGTCACGACTTGGCCCTAGTGTTCGCGCAGATCACCGGCTTGCCTCTT 566
Db      |||
Db      5015 ACCACTGCGGACGAGGCACTGTGCTGCTGCTCCCAACCGGACACCGGCTGCCC----- 4961
QY      567 TTCGCGCAAGGAAATTCGCTGTTGATTAATCTGCGCGGGAAGAACCTTTCGTCGGATGCG 626
Db      |||
Db      4960 ----AGACGCGCGCGGTGGGACCGCGGAGTGGCGGATGAGCGCTTCGTGGCACTGA 4905
QY      627 AGCAGGTTTCGCGCACCGGCTCTCATGATGATGATAGCGGAAGACCGGTTTGTTC 686
Db      |||
Db      4904 GCGGGTCTTCGGAATCCAGTACCGACCGGCTGTGCGCGCGCGGATTCGTGCT 4845
QY      687 CAATGTGGTTCGAAATCCATGGAACTCACCACTGCGAGGCTTGTTCAGCGCAGGCTCT 746
Db      |||
Db      4844 GCGGGTCTTCGGAATCCAGTACCGACCGGCTGTGCGCGCGGCTTCGTGCTGCGCGGCT 4785
QY      747 CGCGGTTGTGTGGTTCC 764
Db      |||
Db      4784 CGGTGTGGCGCTGCGCCC 4767

```

RESULT 7  
ABD13720  
ID ABD13720 standard; DNA; 927 BP.  
XX

```

AC      ABD13720;
XX
DT      29-JUL-2004 (first entry)
XX
DE      Pseudomonas aeruginosa polynucleotide #12324.
XX
KW      Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX      antibacterial.
XX
OS      Pseudomonas aeruginosa.
XX
PN      US5551795-B1.
XX
PD      22-APR-2003.
XX
PF      18-FEB-1999; 99US-00252991.
XX
PR      18-FEB-1999; 98US-0074788P.
XX
PR      27-JUL-1999; 98US-0094190P.
XX
PA      (GENO-) GENOME THERAPEUTICS CORP.
XX
PI      Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
WI      WPI; 2003-615309/58.
XX      P-PSDB; ABO80149.
XX
PT      Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX      useful as molecular targets for diagnostics, prophylaxis and treatment of
XX      pathological conditions resulting from bacterial infection.
XX
PS      Disclosure; SEQ ID NO 12324; 455bp; English.
XX
CC      The invention relates to Pseudomonas aeruginosa polypeptides and the
XX      polynucleotides encoding them. The sequences are useful in diagnosis and
XX      therapy of pathological conditions, as molecular targets for diagnostics,
XX      prophylaxis and treatment of pathological conditions resulting from a
XX      bacterial infection, for evaluating a compound, such as a polypeptide,
XX      for the ability to bind a P. aeruginosa nucleic acid, as components of
XX      effective antibacterial targets, as targets for antibacterial drugs,
XX      including anti-P. aeruginosa drugs, as templates for recombinant
XX      production of P. aeruginosa-derived peptides or polypeptides, as target
XX      components for diagnosis and/or treatment of P. aeruginosa-caused
XX      infection, and in detection of P. aeruginosa sequences or other sequences
XX      of Pseudomonas species using biochip technology. Sequences ABD01397-
XX      ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX      The sequence data for this patent did not form part of the printed
XX      specification but was obtained in electronic format from USPTO at
XX      seqdata.uspto.gov/sequence.html
XX
SQ      Sequence 927 BP; 161 A; 355 C; 279 G; 132 T; 0 U; 0 Other;

```

```

Query Match      7.3%; Score 68; DB 11; Length 927;
Best Local Similarity 49.2%; Pred. No. 2.3e-09;
Matches 179; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY      38  ACGACTACGACGCTTCAATTCAGTCGCTCAATCAGGCCACCTCACCGAAATGCCGAAA 97
Db      |||
Db      44  ACAGCTCAACACCTTCATCGCATTCGCGAGACCGGAGCTTTCGAGAGCGGGGAGC 103
QY      98  GATTAGGATCCCGAGCCACACATTTCCAGACGAATCAGCCGAGTGGGAAACACGACG 157
Db      |||
Db      104  GCCTGACCTGACCCCAACCGGGGTGAGCAAGCGCATCGCCCTCGAACAGCAGCTCA 163
QY      158  GCACCCGCTTTTCAGCGCGCGCGGAACTGCTCTCAACCAACGAGGCGCGCT 217
Db      |||
Db      164  ATGCCCGGCTGTTCGACCGGGTCCGGCGGAGTCAACCTCACCGAGCGCGCGCC 223
QY      218  TCCTCAACACCGCAGCGCCATCGTTCGAGAAATCAACTCCCGCAACTGAAATCAAA 277
Db      |||
Db      224  TGTGCGCGCGGCTTACACGTTGCTCAACGCTGCTGACGACACCCCGCGGGGCTGAC 283
QY      278  GCCTCATGACCCAGAAAAGGCAATCCGACTGAGACTTCATGCTTCCTTTGGGACTT 337

```

```

Db      284 ACCTACCGCGAGGTCACGCGCGGCTGCTCTGGCCACACGACACATCGCGCTGC 343
QY      338 GGATGGTCCCGCACTTATCCGAACAATTCGGCGCGGAGAACACCCCAAGTAGAATTCGAAC 397
Db      344 ATCGCTGCGCGCGCTGCTGCGCGCTTACCGCGCGCATCCGCGAGGTGGCGCTGGATA 403
QY      398 TCCA 401
Db      404 TCCA 407

RESULT 8
ABD13492/c
ID ABD13492 standard; DNA; 972 BP.
XX
AC ABD13492;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polynucleotide #12096.
XX
KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
XX
DR P-PSDB; ABO79921.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 12096; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 972 BP; 141 A; 287 C; 368 G; 176 T; 0 U; 0 Other;
XX
Query Match 7.3%; Score 68; DB 11; Length 972;
Best Local Similarity 49.2%; Pred. No. 2.3e-09;
Matches 179; Conservative 0; Mismatches 185; Indels 0; Gaps 0;
QY 38 ACGACCTACGCGAGCTTCATTTTCAGTCGCTCAATCAGCGCCACCTCACCGAAATCGCGGAA 97

```

```

Db      890 ACAGCTCAACACCTTCATCCCATTCGCGAGACCGGAGCTTCTCCGAACCGCGGAGC 831
QY      98 GATTAGGATCCCGAGCCCACTTTCCAGACGAATCAGCCGAGTGGAAAAACACGCGAG 157
Db      830 GCCTGCACCTGACCAACCGCGCGGTGAGCAAGCGCATCGCGCGCTTCGAACAGAGCTCA 771
QY      158 GCACCCCACTTTTCGACCGCGCGCGGCGCAAACTCGTCTCTCAACGACGAGGCGCAGCCT 217
Db      770 ATGCCCGCTGTTTCGACCGCGGTGCGCGCGGAGGTCAACCTCACCGAGCGCGCGCGCCC 711
QY      218 TCCTCAACACGCGCGCGCATTCGTCGAGAAATCAACTCCGCGCGCAACTGAAATCAAAC 277
Db      710 TGCTCGCGCGCTTACCAGTTGCTCAAGTGTCTGACGACACCGCGCGGCGGTGAACA 651
QY      278 GCCTCATGAGCCCAAGAAAAGGCAATCCGACTTGGACTTTCATGCAATTCCTTGGGCACTT 337
Db      650 ACCTCACCGCGAGGTGAGCGCGCGGTGCTCTGGCCACCGACCAATCGCGCTGC 591
QY      338 GGATGGTCCCGAACTTATCCGACATTCGCGCGCGGACACCCCAACGTAAGATCCCAAC 397
Db      590 ATCGCTGCGCGCGCTGCTGCGCGCTTCACCGCGCGCATCCGAGGTGGCGTGGATA 531
QY      398 TCCA 401
Db      530 TCCA 527

RESULT 9
ABD13817
ID ABD13817 standard; DNA; 1482 BP.
XX
AC ABD13817;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polynucleotide #12421.
XX
KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
XX 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
XX
DR P-PSDB; ABO80246.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 12421; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant

```

CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABD01397-  
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 1482 BP; 269 A; 555 C; 431 G; 227 T; 0 U; 0 Other;  
Query Match 7.3%; Score 68; DB 11; Length 1482;  
Best Local Similarity 49.2%; Pred. No. 2.7e-09;  
Matches 179; Conservative 0; Mismatches 185; Indels 0; Gaps 0;  
QY 38 ACGAGCTACGCGAGCTTCATTTCAGTGCCTCAATCAGGCGCACCTCACCAGAACTGCCGAAA 97  
DB 99 ACAGCTCAACACTTCATCGCCATTGCGAGACCGGAGCTTCTCCGAGCGGGAGC 158  
QY 98 GATTAGGATCCCGCAGCCACACACTTTCAGAGCAATCAGCGAGTGGAAAAACACGCGAG 157  
DB 159 GCCTGCACCTGACCCAAACCGCGGTGAGCAAGCGCATCGCCCTCGAACACGAGCTCA 218  
QY 158 GCACCCCACTTTGACCGCGCGCGGCGGCAAACTGCTCTCAACCAACGAGGCGCGCT 217  
DB 219 ATGCCCGGCTGTTCGACCGGGTCCGCGCGAGGTCAACCTCACCAGGCGCGCGCGCC 278  
QY 218 TCTCAACACGCGCGCGCTTCGCGAGATTCAGCTCCGCGCACTGAAATCAAC 277  
DB 279 TGTGCGCGCGCTTACCAAGTTGCTCAACGTCTCGACGACACCGCGCGGCTGAACA 338  
QY 278 GCTCATGACCGCAGAAAAAGGACAAATCCGACTGGACTTCATGATTCCTTTGGGCACTT 337  
DB 339 ACCTACCGCGGAGGTACGCGCGCGGTGCTCTGCGCACAGCCACACATCGGCTGC 398  
QY 338 GGATGCTCCCGAATTATCCGAACATTCGCGCGCGAAACACCCCAACGTTAGAAATCCCAAC 397  
DB 399 ATGCGCTGCGCGCTGTGCTGCGCGCTTCAACCGCGCGCATCCGCGAGTGGCGCTGGATA 458  
QY 398 TCCA 401  
DB 459 TCCA 462

RESULT 10  
ABD08478  
ID ABD08478 standard; DNA; 987 BP.  
AC ABD08478;  
XX  
XX 29-JUL-2004 (first entry)  
XX  
XX Pseudomonas aeruginosa polynucleotide #7082.  
DE Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;  
KW antibacterial.  
XX  
XX Pseudomonas aeruginosa.  
XX  
XX US6551795-B1.  
XX  
XX 22-APR-2003.  
XX  
XX 18-FEB-1999; 99US-00252991.  
XX  
XX 18-FEB-1998; 98US-0074788P.  
XX  
XX 27-JUL-1998; 98US-0094190P.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX WPI; 2003-615309/58.  
XX

DR P-PSDB; AB074907.  
XX  
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, if  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
XX Disclosure; SEQ ID NO 7082; 455pp; English.  
XX  
CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABD01397-  
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 987 BP; 189 A; 352 C; 298 G; 148 T; 0 U; 0 Other;  
Query Match 6.6%; Score 61.4; DB 11; Length 987;  
Best Local Similarity 47.7%; Pred. No. 2.1e-07;  
Matches 179; Conservative 0; Mismatches 196; Indels 0; Gaps 0;  
QY 27 COTCGGAATCGACGACCTACGCGAGCTTCATTTTCAGTGCCTCAATCAGGCGCACCTCACCGA 86  
DB 78 CTTGGCTTTCGAGGATATCGAAGCTTCTCAACGTCGCGGAATCAACAGGCTGAGGTC 137  
QY 87 AACTCCGGAAGATTAGGATCCCGCAGCCCACTTTTCAGACGAATCAGCGGAGTGGA 146  
DB 138 GCGCGCGGAGCGCATGAGCTGTCCAAGTGGTGATCAGCAAGCGTGTGAGGATCTCGA 197  
QY 147 AAACACGCGGCGACCCCACTTTTCGACGCGCGCGCGCAAACTCGTCTCAACCAACG 206  
DB 198 ACGAGCTCTGGCGGTCCGCTCTATCGACACGCGCGCAACGTCGAGCCGACGAGGC 257  
QY 207 AGGCCACGCTTCTCTCAACACGCGCGCGGCGGCACTCGTCGAGAAATCAATCCGCGCAAC 266  
DB 258 CGGCAATTTCTTCTACAGGCGCGCAAGGCTCGCTGCGAGGAATCAACACGCGCGGA 317  
QY 267 TGAATCAAAAGCCTCATGAGCCCAAGAAAAGGCAATCCGACTGGACTTCATGATTC 326  
DB 318 GATTTCGCGCTCGCGGAGAACGACCTGTGCGCGGAATTCGCGCATCGTCACCCCGATGAG 377  
QY 327 CTTGGGCACTTGGATGTTCCCGGAATTCGGAACATTCGCGCGGCGGACACCCCAACCT 386  
DB 378 CTTGCGCACCTCTGCTGTGCGGATCATCGCGGACTTCATGAGCCAGCACCCGCGCT 437  
QY 387 AGAATTCCAACTCCA 401  
DB 438 GGAGATCGTCTGCA 452  
RESULT 11  
ABD08454  
ID ABD08454 standard; DNA; 1485 BP.  
XX  
XX ABD08454;  
XX  
XX 29-JUL-2004 (first entry)  
XX  
XX Pseudomonas aeruginosa polynucleotide #7058.  
XX  
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;  
XX antibacterial.  
XX



OS Pseudomonas aeruginosa.  
XX US6551795-B1.  
XX 22-APR-2003.  
XX 18-FEB-1999; 99US-00252991.  
XX 18-FEB-1998; 98US-0074788P.  
XX 27-JUL-1998; 98US-0094190P.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX WPI; 2003-615309/58.  
XX P-PSDB; ABO74883.  
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX pathological conditions resulting from bacterial infection.  
XX Disclosure; SEQ ID NO 7058; 455pp; English.  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using biochip technology. Sequences ABD01397-  
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:  
XX The sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html  
XX  
XX Sequence 1485 BP; 297 A; 531 C; 443 G; 214 T; 0 U; 0 Other;  
Query Match 6.6%; Score 61.4; DB 11; Length 1485;  
Best Local Similarity 47.7%; Pred. No. 2.5e-07;  
Matches 179; Conservative 0; Mismatches 196; Indels 0; Gaps 0;  
QY 27 CCTGCGAATCGACGACCTACGAGCTTTCATTTTCAGTCGCTCAATCAGGCCACCTCACCGA 86  
DB 124 CTGCGTTTCGACGATATCGAGCTTCTCAACGTCGCGNACTCAACAGGTGACGTC 183  
QY 87 AACTGCCGAAGATTAGGATCCCGAGCCACACTTTTCAGACGAATCAGCCGAGTGGGA 146  
DB 184 GCGCGCCGAGCGCATGAGCTGTCCAAGTCGCTGATCAGCAAGCTGTGACGATCTCGA 243  
QY 147 AAAACACGACGACGACCCACATTTTCGACGCGCGCGCGCAACTCGTCTCAACCAACG 206  
DB 244 ACGAGCTCTGCGCGTCCGCTCTATCGACACGCGCGCACTGCGAGCGCGACGAGGC 303  
QY 207 AGGCCACGCGCTTCTCAACACGCGCGCGCCATCGTCGAGAAATTCACACTCGCGCAAC 266  
DB 304 CGGCGATTCTTCTACAGGCGCGCAAGGCTCGTGCAGGNACTCAACAGCGCGCGGA 363  
QY 267 TGAATCAACGCGCTATGGACCCAGAAAAGGCAATTCGACATGGAATTCATGCAATC 326  
DB 364 GATTTCGCGCTCGCGGAGAACGACCTGTGCGCGCAATTTGGCATCTGTCACCCCGATGAG 423  
QY 327 CTTGGGCACTTGGATGTTCCCGAATTTATCGACATTTCCGCGCGCAACCCCAAGT 386  
DB 424 CTTCCGCACTTCTGCTGTGCTCCGATCATCGCGCACTTCATGAGCCAGCACCGCGCT 483  
QY 387 AGAATTCCAACTCCA 401

DB 484 GGAGATCGTCTCTGCA 498  
RESULT 12  
ABD08440/c  
XX ABD08440 standard; DNA; 1518 BP.  
XX AC ABD08440;  
XX DT 29-JUL-2004 (first entry)  
XX Pseudomonas aeruginosa polynucleotide #7044.  
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;  
XX antibacterial.  
XX Pseudomonas aeruginosa.  
XX US6551795-B1.  
XX 22-APR-2003.  
XX 18-FEB-1999; 99US-00252991.  
XX 18-FEB-1998; 98US-0074788P.  
XX 27-JUL-1998; 98US-0094190P.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX WPI; 2003-615309/58.  
XX P-PSDB; ABO74869.  
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX pathological conditions resulting from bacterial infection.  
XX Disclosure; SEQ ID NO 7044; 455pp; English.  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using biochip technology. Sequences ABD01397-  
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:  
XX The sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html  
XX  
XX Sequence 1518 BP; 213 A; 451 C; 554 G; 300 T; 0 U; 0 Other;  
Query Match 6.6%; Score 61.4; DB 11; Length 1518;  
Best Local Similarity 47.7%; Pred. No. 2.5e-07;  
Matches 179; Conservative 0; Mismatches 196; Indels 0; Gaps 0;  
QY 27 CCTGCGAATCGACGACCTACGAGCTTTCATTTTCAGTCGCTCAATCAGGCCACCTCACCGA 86  
DB 1429 CCTCGTTTCGACGATATCGAAGCTTCTCAACGTCGCGNACTCAACAGGTGACGTC 1370  
QY 87 AACTGCCGAAGATTAGGATCCCGAGCCACACTTTTCAGACGAATCAGCCGAGTGGGA 146  
DB 1369 GCGCGCCGAGCGCATGAGCTGTCCAAGTCGCTGATCAGCAAGCTGTGACGATCTCGA 1310  
QY 147 AAAACACGACGACGACCCACATTTTCGACGCGCGCGCGCAAACTCGTCTCAACCAACG 206

Db 1309 ACAGCGCTCGGCTCGCTCTATCCACACCGCGCAACGTCGAGCCGACGAGGC 1250  
QY 207 AGCCACGCTTCTCAACACCGCCAGCGCATCGCGAGAAATTCAAATTCGCCGCCCAAC 366  
Db 1249 CGGGCATTTCTTACAGGCGCCCAAGGCTCGCTGCAGGAATCAACACGCGCGGA 1190  
QY 267 TGAATCAACGCTCATGACCCAGAAAGGCAATCCGACTGAGCTTCATGATTC 326  
Db 1189 GATTGTGCGCTCGCGGAGAACACCTGTGCGCGGAATTGCGCATGCTCACCCCGATGAG 1130  
QY 327 CTGGGCACTTGGATGCTCCCGAACTTATCCGAACATTCGCGCGGCAACCCCAACCT 386  
Db 1129 CTTCGCAACCTCTGCTGTCGCGGATCATGCGCGACTTCATGAGCAGACCCGCGCT 1070  
QY 387 AGAATTCCAACTCCA 401  
Db 1069 GGAGATCGTCTGCA 1055

## RESULT 13

ACA43570

ID ACA43570 standard; DNA; 891 BP.

AC ACA43570;

XX

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #25227.

XX

KW Antisense; ds; prokaryotic essential gene; cell proliferation;

KW drug design; gene.

XX

OS Pseudomonas putida.

XX

XX WO20027183-A2.

XX

XX 03-OCT-2002.

XX

XX 21-MAR-2002; 2002WO-US009107.

XX

XX 21-MAR-2001; 2001US-00815242.

XX

XX 06-SEP-2001; 2001US-00948993.

XX

XX 25-OCT-2001; 2001US-0342923P.

XX

XX 08-FEB-2002; 2002US-00072851.

XX

XX 06-MAR-2002; 2002US-0362699P.

XX

XX (ELIT-) ELITRA PHARM INC.

XX

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen K, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

XX WPI; 2003-029926/02.

XX

XX P-PSDB; ABU39700.

XX

PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.

XX

PS Claim 14; SEQ ID NO 31440; 1766pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 891 BP; 163 A; 294 C; 270 G; 164 T; 0 U; 0 Other;

Query Match 6.5%;

Best Local Similarity 46.4%;

Pred. No. 5.2e-07;

Matches 195; Conservative 0;

Mismatches 225; Indels 0;

Gaps 0;

QY 38 ACGACTACGACGCTTCAATTCAGTCGCTCAATCAGGCCACCTCACCGAAATGCCGAAA 97

Db 17 AGACATCAGATCTTTCGTGAATAATCGCGAATCTGGGAGTTTCATGAGGCGCGCAAG 76

QY 98 GATTAGGCATCCCGCAGCCCACTTTCCAGACGAATCAGCCGAGTGGAAAAACAGCAG 157

Db 77 CGCTCCACTTGTCCAGCGCGCCCTGAGCCGACGATCAAGAAGCTGGAGGAGGCTCG 136

QY 158 GCACCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACGAGGCCAGCCT 217

Db 137 GCACCTCGCTACTTGAACGCACTCTCGCGGCTCAGCTGAGCGGCGCGGCT 196

QY 218 TCCTCAACACGCGCAGCGCATCTGTCGAGAAATTCACCTCCGCGCACTGAAATCAAC 277

Db 197 TCCTGCCAAGCCAGCGCGCTGCTGGATGATTTTGAAGACTCGATCTCAGCATCCGCG 256

QY 278 GCCTCATGACCCAGAAAAAGSCACAATCCGACTGGACTTCATGATTCCTTTGGGCACTT 337

Db 257 AGCTGGCGAGCGCCAGACCGGTACGTTACCTCGCTGCAATCTCTACCGCGGCTTCT 316

QY 338 GGATGGTCCCGAATTCGGAACATTCGCGCGCGCAACACCCCAACGTAGAAATCCCAAC 397

Db 317 ACTTCTCGCTGCGGTGATTCGCGGATTAACAGAGCAATACCCGAAATTCGCAATTCGCC 376

QY 398 TCCACCAAGCGCAGCAATGCTCCTGGTAGATCGTGTGCTGATGAACACTGACCTCG 457

Db 377 TGCTGGACCTTAGCGCCCAACGCGGCTCGAAGCCGCTGCTGCGCGGAGGCGGCTTGG 436

## RESULT 14

AAH68525/c

ID AAH68525 standard; DNA; 349980 BP.

XX AC AAH68525;

XX

XX 26-SEP-2001 (first entry)

XX

XX C glutamicum coding sequence fragment SEQ ID NO: 7060.

XX

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;

XX

XX organic acid synthesis; ds.

XX

XX Corynebacterium glutamicum.

XX

XX EP1108790-A2.

XX

XX 20-JUN-2001.

XX

XX

PF 18-DEC-2000; 2000EP-00127698.  
XX  
PR 16-DEC-1999; 93JP-00377484.  
PR 07-APR-2000; 2000JP-00159182.  
XX 03-AUG-2000; 2000JP-00280988.  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX WPI; 2001-376931/40.  
XX  
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analyzing  
PT expression profile or pattern of a gene and identifying homologous gene.  
XX  
XX Disclosure; SEQ ID NO 7060; 246pp + Sequence Listing; English.  
XX  
XX The present invention provides a number of nucleotide and protein  
CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of corynebacterium, measuring expression amount and analyzing  
CC the expression profile or expression pattern of a gene derived from  
CC Corynebacterium, and identifying a homologue of a gene derived from  
CC Corynebacterium. Corynebacterium bacteria are useful for producing amino  
CC acids, nucleic acids, vitamins, saccharides and organic acids, described  
CC particularly L-lysine. The present sequence is a nucleic acid, described  
CC in the exemplification of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the European Patent Office  
XX  
SQ Sequence 349980 BP; 80289 A; 91081 C; 97378 G; 81232 T; 0 U; 0 Other;  
Query Match 6.5%; Score 60; DB 5; Length 349980;  
Best Local Similarity 51.5%; Pred. No. 6.2e-06;  
Matches 138; Conservative 0; Mismatches 130; Indels 0; Gaps 0;  
QY 22 GGAGACCTGGGAATCGACGACCTACGACGCTTCATTCAGTCGCTCAATCAGCGCACCTC 81  
Db 142488 GCACACATGCACTCAATCAGCTCGAATTTTCATCGCAGTAGGCCAACACGGACGATC 142429  
QY 82 ACCGAAATCGCGAAGATGAGCATCCGCGAGCCGACACTTTCAGACGCAATCAGCGCA 141  
Db 142428 AACCGCGCGCGAGAACTCTCTCATTTCCCAACCGCTCTCAGCGCAGATCTCCGCA 142369  
QY 142 GTGGAATAACAGCAGGACCCACTTTTCAGACGCGCGCGGCGCAAACTCGTCTCTAAC 201  
Db 142368 CTTGAAAAATCGTCGGAGCTCCACTTTCGAAACGCCATTCGCGGGTGTCTCCCTACA 142309  
QY 202 CAACGAGGCGCGCTTCTCAACACGCGCGCATCGTCGCGAATTCACCTCCGCC 261  
Db 142308 AAGCGCGGAGAAATCTTCAGGAAGAGCCCTCCGAAGCTTAGCAGGATGCAATCGGTA 142249  
QY 262 GCAACTGAAATCAACCGCTCATGACC 289  
Db 142248 GTCGATGAAATCAATCCGGTAGGACC 142221

RESULT 15  
ACA26845  
ID ACA26845 standard; DNA; 954 BP.  
XX  
AC ACA26845;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Prokaryotic essential gene #8502.  
XX  
KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
XX drug design; gene.  
OS Bordetella pertussis.

XX WO200277183-A2.  
PN  
XX 03-OCT-2002.  
PD  
XX 21-MAR-2002; 2002WO-US009107.  
PF  
XX 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX (ELIT-) ELITRA PHARM INC.  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-022926/02.  
DR P-PSDB; ABU22975.  
DR  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 14; SEQ ID NO 14715; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 954 BP; 149 A; 350 C; 302 G; 153 T; 0 U; 0 Other;  
Query Match 6.4%; Score 59.8; DB 8; Length 954;  
Best Local Similarity 47.5%; Pred. No. 6.1e-07;  
Matches 178; Conservative 0; Mismatches 197; Indels 0; Gaps 0;  
QY 25 GACCTGGCAATCGACGACCTAGCGAGCTTCAATTCAGTCGCTCAATCAGGCGACCTCACC 84  
Db 16 GATCTGTCCACCCCGCAATTCGCGGCTTCGCTCATGCTTGGCGACCTCGCGACCTTACC 75  
QY 85 GAAACTGCCGAAGATTAGGCATCCGCGAGCCCACTTTCCAGACGCAATCAGCGGAGTG 144  
Db 76 AAGCGCGCGAGGATGCGCATTTGTCCAGCCCGCATTCAGGCGCACTGATCCAGTCCCTG 135



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2005, 23:13:05 ; Search time 3101 Seconds  
(without alignments)

11415.598 Million cell updates/sec

Title: US-09-826-909-1\_COPY\_232\_1161

Perfect score: 930

Sequence: 1 gtagcattgggcaacgacgg.....ggtagtcattagaagagggc 930

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsa1:\*

9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 67.4  | 7.2         | 760    | 9     | CC842306    |
| 2          | 57.8  | 6.2         | 766    | 8     | BZ573361    |
| 3          | 57.8  | 6.2         | 847    | 8     | BZ553158    |
| 4          | 57.8  | 6.2         | 1350   | 8     | BZ554839    |
| 5          | 57.6  | 6.2         | 1175   | 8     | BZ568637    |
| 6          | 57.4  | 6.2         | 990    | 8     | BZ573643    |
| 7          | 56.8  | 6.1         | 1077   | 8     | BZ563333    |
| 8          | 56.2  | 6.0         | 717    | 7     | CN822018    |
| 9          | 56.2  | 6.0         | 768    | 7     | CN822012    |
| 10         | 56.2  | 6.0         | 853    | 7     | CO202977    |
| 11         | 55.8  | 6.0         | 850    | 8     | BZ424306    |
| 12         | 55.6  | 6.0         | 817    | 8     | BZ566287    |
| 13         | 55.2  | 5.9         | 674    | 8     | AZ933780    |
| 14         | 55    | 5.9         | 1072   | 8     | BZ579898    |
| 15         | 54.8  | 5.9         | 451    | 8     | BZ422864    |
| 16         | 54.6  | 5.9         | 800    | 7     | CF933278    |
| 17         | 54.4  | 5.8         | 1161   | 8     | BZ556936    |
| 18         | 54.2  | 5.8         | 391    | 9     | P294R       |
| 19         | 53.4  | 5.7         | 780    | 8     | BZ560257    |
| 20         | 53.2  | 5.7         | 1106   | 8     | BZ559233    |
| 21         | 53    | 5.7         | 722    | 8     | AZ933651    |
| 22         | 52.8  | 5.7         | 551    | 1     | A1409183    |
| 23         | 52    | 5.6         | 815    | 7     | CN822455    |
| 24         | 52    | 5.6         | 820    | 7     | CO202387    |

|    |      |     |      |   |           |            |
|----|------|-----|------|---|-----------|------------|
| 25 | 51   | 5.5 | 721  | 8 | BZ562255  | BZ562255   |
| 26 | 50.4 | 5.4 | 729  | 7 | CN822500  | CN822500   |
| 27 | 50.2 | 5.4 | 331  | 2 | BF291112  | BF291112   |
| 28 | 50.2 | 5.4 | 806  | 8 | BZ551243  | BZ551243   |
| 29 | 50.2 | 5.4 | 1291 | 8 | BZ568924  | BZ568924   |
| 30 | 50   | 5.4 | 734  | 8 | AY080301  | AY080301   |
| 31 | 49.2 | 5.3 | 1013 | 8 | BZ548324  | BZ548324   |
| 32 | 48.8 | 5.2 | 1212 | 8 | BZ579976  | BZ579976   |
| 33 | 47.6 | 5.1 | 673  | 8 | BZ548565  | BZ548565   |
| 34 | 47.4 | 5.1 | 1327 | 8 | BZ571615  | BZ571615   |
| 35 | 47   | 5.1 | 406  | 7 | CO000903  | CO000903   |
| 36 | 46.6 | 5.0 | 1280 | 8 | BZ578104  | BZ578104   |
| 37 | 46.2 | 5.0 | 636  | 8 | BH190115  | BH190115   |
| 38 | 46.2 | 5.0 | 922  | 9 | CNS0073W  | AL066784   |
| 39 | 46.2 | 5.0 | 1512 | 8 | BZ557409  | Drosophila |
| 40 | 46.2 | 5.0 | 2328 | 9 | CL982230  | CL982230   |
| 41 | 46   | 4.9 | 884  | 9 | CNS006U0  | CL055923   |
| 42 | 46   | 4.9 | 909  | 9 | CNS000JTL | Drosophila |
| 43 | 46   | 4.9 | 1052 | 8 | BZ559123  | AL076720   |
| 44 | 45.4 | 4.9 | 371  | 7 | CNS586944 | Drosophila |
| 45 | 45.4 | 4.9 | 861  | 8 | BZ572433  | CN586944   |

ALIGNMENTS

RESULT 1  
CC842306  
LOCUS ND.L105A23.T7 Notre Dame Liverpool Aedes aegypti genomic clone  
DEFINITION CC842306  
ACCESSION CC842306  
VERSION CC842306.1 GI:33188999  
KEYWORDS GSS:  
SOURCE Aedes aegypti (yellow fever mosquito)  
ORGANISM Aedes aegypti  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;  
Stegomyia.

REFERENCE 1 (bases 1 to 760)  
Lofthus,B., Shetty,J., Knudson,D. and Severson,D.  
BAC end sequencing of Aedes aegypti  
Unpublished (2003)  
JOURNAL  
COMMENT Other GSSs: ND.L105A23.SP6  
Contact: Brendan Loftus  
Department of Eukaryotic Genomics

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208  
Email: enta@tigr.org  
Library was provided by David Severson  
Seq primer: 17  
Class: BAC ends.

FEATURES  
Location/Qualifiers  
1..760  
/organism="Aedes aegypti"  
/mol\_type="genomic DNA"  
/strain="Liverpool"  
/db\_xref="taxon:7159"  
/clone="Notre Dame Liverpool-105A23"  
/clone.lib="Notre Dame Liverpool"  
/note="Vector: pECBAC1; Site\_1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN

Query Match 7.2%; Score 67.4; DB 9; Length 760;  
Best Local Similarity 50.4%; Pred.No. 5.9e-08;  
Matches 192; Conservative 0; Mismatches 186; Indels 3; Gaps 1;  
Qy 78 CCTCACCAGAACTGCCGAAAGATTAGGCATCCCGCAGCCACACTTTCAGCAATCAG 137

Db 3 CTTTACCCCGCGGGGAAAAAATCTTTCGTCAACCCAGCCACCATCAGCAAGATGCTCGG 62  
 Qy 138 CCGAGTGAATAACAGCAGGACCCCACTTTTCGACCGCGCGCCGCCCAACTGCTCT 197  
 Db 63 GCATCTGGAAGAGGAGCTGGAATGCACGCTGTGATCCGCGAAGCGCAAGCTGCGCT 122  
 Qy 198 CAACCAACGAGCCAGCGCTTCTCAACACGCGAGCGCCATCGTCGAGAAATCAACTC 257  
 Db 123 GACCGACAGCGCCAGCGGCTGTATCAGCGGCGCTGACGATCTCGAGAAATCCGCA 182  
 Qy 258 CGCCGCAACTGAAATCAACGCGCTCATGAGCCAGAAAAGGACCAATCCGACTGACTT 317  
 Db 183 GCTGAGCGGAGCTGGAAGACATCAGCTCGTGAAAAAGGCGTGCTGGCGTGGCAT 242  
 Qy 318 CATGATCTCTTGGCACTTGATGCTCCCGAATTCGGAATTCGCGCGCAACA 377  
 Db 243 TCCGCGGATGTTGGGCA---GCGAGATCGCGATCTGATCGCGCTTTTCGCCAGACCTA 299  
 Qy 378 CCCCAACGTAGAATTCCAACTCCACCAAGCGCAGCAATGCTCTGGTAGATCGGTGTTT 437  
 Db 300 CCCCGGATCGAATTAATAATTCGAGCTGGCGGCTGTGCTGCAAGCGGCTGAT 359  
 Qy 438 GGCTGATGAACCTGACCTCGC 458  
 Db 360 GTCCGCGAAGCTGGATCTGGC 380

RESULT 2  
 BZ573361  
 LOCUS msh2\_307.x3 msh Pseudomonas aeruginosa genomic clone msh2\_307, 766 bp DNA linear GSS 17-DEC-2002  
 DEFINITION genomic survey sequence.

ACCESSION BZ573361  
 VERSION BZ573361.1 GI:27208422  
 KEYWORDS GSS.

SOURCE Pseudomonas aeruginosa  
 ORGANISM Pseudomonas aeruginosa  
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 766)  
 AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.  
 TITLE Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library  
 JOURNAL J. Bacteriol. (2002) In press  
 COMMENT Contact: Chris K. Raymond  
 Genome Center  
 University of Washington  
 Box 352145, Seattle, WA 98105-2145, USA  
 Tel: 2062216954  
 Fax: 2066857244  
 Email: craymond@u.washington.edu  
 Class: shotgun.

FEATURES  
 source  
 Location/Qualifiers  
 1..766  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="MSH"  
 /db\_xref="taxon:287"  
 /clone="msh2\_307"  
 /clone\_lib="msh"  
 /note="Environmental isolate. Whole genomic shotgun library."

ORIGIN  
 Query Match 6.2%; Score 57.8; DB 8; Length 766;  
 Best Local Similarity 52.7%; Pred. No. 3.4e-05;  
 Matches 125; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 27 CTTGGAATCGACGACCTAGCAGCTTCATTTTCAGTCGCTCATCAGGCGACCTCACCGA 86  
 Db 301 CTTGCGCTTTGAAGACATACAGGCTTCTCTCAACGTGGTCGAGCTGGGCGATCAGCGC 360

Qy 87 AACTGCCGAAAGATTAGGCATCCCGCAGCCACACTTTTCCAGACGAATCAGCCGAGTGA 146  
 Db 361 CGCGCGGAGCGATGAGCTGTCAAGTCGTGATCAGCAACGCGTCAGGACCTGA 420  
 Qy 147 AAAACAGCAGCAGGACCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACG 206  
 Db 421 CGCCACCTCGCGTGGCGCTCTACCGCTCGACGCGCAACGTCGAGCGCGACCGAGGC 480  
 Qy 207 AGGCACGCTTCTCTCAACACGCGCGGCATCGTCGAGAAATTCAACTCGCGGC 263  
 Db 481 CGCGGCTTCTTCTACAAAGTCGCGCAAGGCTCTCGTCGAGACCTGAAACAACGCGCG 537

## RESULT 3

BZ553158/c

LOCUS

DEFINITION BZ553158 847 bp DNA linear GSS 17-DEC-2002  
 pacsl-60\_4007.y2 pacsl-60 Pseudomonas aeruginosa genomic clone  
 pacsl-60\_4007, genomic survey sequence.

ACCESSION BZ553158

VERSION BZ553158.1

KEYWORDS GSS.

SOURCE Pseudomonas aeruginosa

ORGANISM Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 847)

AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.

TITLE Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library

JOURNAL J. Bacteriol. (2002) In press

COMMENT Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

## FEATURES

source

Location/Qualifiers  
 1..847  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="1-60"  
 /db\_xref="taxon:287"  
 /clone="pacsl-60\_4007"  
 /clone\_lib="pacsl-60"  
 /note="clinical isolate 1-60 Whole genomic shotgun library."

## ORIGIN

Query Match 6.2%; Score 57.8; DB 8; Length 847;  
 Best Local Similarity 52.7%; Pred. No. 3.4e-05;  
 Matches 125; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 27 CTTGGAATCGACGACCTAGCAGCTTCATTTTCAGTCGCTCATCAGGCGACCTCACCGA 86  
 Db 408 CTTGCGCTTTGAAGACATACAGGCTTCTCTCAACGTGGTCGAGCTGGGCGATCAGCGC 349

Qy 87 AACTGCCGAAAGATTAGGCATCCCGCAGCCACACTTTTCCAGAGAAATCAGCCGAGTGA 146  
 Db 348 CGCGCGGAGCGCATGAGCTGTCCAAGTCGTGATCAGCAACGCGTCAGGACCTGA 289

Qy 147 AAAACAGCAGGACCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACG 206  
 Db 288 CGCGCACCTCGCGGTGGCGCTCTACCGCTCGACGCGCAACGTCGAGCGCGACGAGGC 229

Qy 207 AGGCACGCTTCTCTCAACACGCGCGGCATCGTCGAGAAATTCAACTCGCGGC 263  
 Db 228 CGCGGCTTCTTCTACAAAGTCGCGCAAGGCTCTCGTCGAGACCTGAAACAACGCGGC 172



```

RESULT 4
BZ554839
LOCUS
DEFINITION
pacsl-60_4884.x1 pacsl-60 Pseudomonas aeruginosa genomic clone
pacsl-60_4884, genomic survey sequence.
ACCESSION
BZ554839.1 GI:27162898
VERSION
BZ554839.1
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 1350)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
JOURNAL
Pseudomonas aeruginosa library
COMMENT
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..1350
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacsl-60_4884"
/clone_lib="pacsl-60"
/note="clinical isolate 1-60 Whole genomic shotgun
library."
FEATURES
source
1..1350
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacsl-60_4884"
/clone_lib="pacsl-60"
/note="clinical isolate 1-60 Whole genomic shotgun
library."
ORIGIN
Query Match 6.2%; Score 57.8; DB 8; Length 1350;
Best Local Similarity 52.7%; Pred. No. 3.8e-05;
Matches 125; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 27 CTTGCGAATCGACGACCTACGAGCTTCATTTTCAGTCGCTCAATCAGGCACTCACCAGA 86
|||||
Db 242 CTTGCGCTTGAAGACATACAGCCTTCTCAACGTGTCGAGCTGGCGAGCATCAGCGC 301
QY 87 AACTCGCGAAGATTAGGATCCCGAGCCACACTTTCCAGACGAATCAGCGGAGTGA 146
|||||
Db 302 CGCCGCCGAACGATGAGCTGTCCAAGTCCGTGATCAGCAACGCGTCAGCGACCTGA 361
QY 147 AAACACGACGACGACCCCACTTTTCGACGCGCGCGCAAACTCGTCTCAACCAAG 206
|||||
Db 362 GCGGCACCTCGCGGTGGCGCTCTTACCGCTCGAGCGCAACGTCGAGCCGACGAGGC 421
QY 207 AGGCCACGCTTCTCTCAACACGACGCGCCACTCGTCGAGAAATCAACTCGCGCGC 263
|||||
Db 422 CGCGCGCTTCTTCTACAGTCGCGCAAGCGCTCGTCGAGGACCTGCAACACGCGCG 478

RESULT 5
BZ568637/c
LOCUS
DEFINITION
pacsl-60_4884.x1 pacsl-60 Pseudomonas aeruginosa genomic clone
pacsl-60_4884, genomic survey sequence.
ACCESSION
BZ568637.1 GI:27201914
VERSION
BZ568637.1
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 1175)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
JOURNAL
Pseudomonas aeruginosa library
COMMENT
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..1175
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_766"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."
FEATURES
source
1..1175
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_766"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."
ORIGIN
Query Match 6.2%; Score 57.6; DB 8; Length 1175;
Best Local Similarity 49.2%; Pred. No. 4.2e-05;
Matches 181; Conservative 0; Mismatches 184; Indels 3; Gaps 1;
QY 39 CGACCTACGACGCTTCATTTTCAGTCGCTCAATCAGGCACTCAGCGAACTCCGAAAG 98
|||||
Db 516 CAGTTCCGCCAGTTTCGTGCGGTTTGGCGAGGAGTGAGCTTCCGCGCGCTGAGCG 457
QY 99 ATTAGGATCCGCGACGCCACACTTTCCAGACGAATCAGCCGAGTGGAAAAACACGACG 158
|||||
Db 456 CTTGACATGGCCAGCCACCGCTGACCTGCGGATCGCGCGCTGGAAGAGAGTCCG 397
QY 159 CACCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACGAGGCCGCGCTT 218
|||||
Db 396 CGCAGCGCTGTGGAGCGTGACACCGTATCCACCGCTGACGCGGCGCGCGCGCTT 337
QY 219 CCTCAACACGCGCGCGCATGTCGCGAGATTTCAACTCCGCGCGCACTGAATCAAACG 278
|||||
Db 336 CTTGACGAAGCGCGCGCGCACCTTGGCGCGCGCGCAACGACCTTGGCGG---CGCGCG 280
QY 279 CTTGATGACCCAGAAAAAGGACAAATCCGACTGACTTTCATGCTTCTTGGGCACTTG 338
|||||
Db 279 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 220
QY 339 GATGTCCTCCGCAACTTATCCGAACTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 398
|||||
Db 219 TCTGCTACGCGCATACTCAGAGCGTTCCGCGCGCGCGCGCGCGCGCGCGCGCG 160
QY 399 CCACCAAG 406
|||||
Db 159 CCAGGAAG 152

RESULT 6
BZ573643
LOCUS
DEFINITION
msh2_3292.x1 msh Pseudomonas aeruginosa genomic clone msh2_3292,
genomic survey sequence.
ACCESSION
BZ573643
VERSION
BZ573643.1 GI:27208704
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 990)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
JOURNAL
Pseudomonas aeruginosa library
COMMENT
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..990
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_766"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."
FEATURES
source
1..990
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_766"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."
ORIGIN
Query Match 6.2%; Score 57.6; DB 8; Length 1175;
Best Local Similarity 49.2%; Pred. No. 4.2e-05;
Matches 181; Conservative 0; Mismatches 184; Indels 3; Gaps 1;
QY 39 CGACCTACGACGCTTCATTTTCAGTCGCTCAATCAGGCACTCAGCGAACTCCGAAAG 98
|||||
Db 516 CAGTTCCGCCAGTTTCGTGCGGTTTGGCGAGGAGTGAGCTTCCGCGCGCTGAGCG 457
QY 99 ATTAGGATCCGCGACGCCACACTTTCCAGACGAATCAGCCGAGTGGAAAAACACGACG 158
|||||
Db 456 CTTGACATGGCCAGCCACCGCTGACCTGCGGATCGCGCGCTGGAAGAGAGTCCG 397
QY 159 CACCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACGAGGCCGCGCTT 218
|||||
Db 396 CGCAGCGCTGTGGAGCGTGACACCGTATCCACCGCTGACGCGGCGCGCGCGCTT 337
QY 219 CCTCAACACGCGCGCGCATGTCGCGAGATTTCAACTCCGCGCGCACTGAATCAAACG 278
|||||
Db 336 CTTGACGAAGCGCGCGCGCACCTTGGCGCGCGCGCGCAACGACCTTGGCGG---CGCGCG 280
QY 279 CTTGATGACCCAGAAAAAGGACAAATCCGACTGACTTTCATGCTTCTTGGGCACTTG 338
|||||
Db 279 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 220
QY 339 GATGTCCTCCGCAACTTATCCGAACTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 398
|||||
Db 219 TCTGCTACGCGCATACTCAGAGCGTTCCGCGCGCGCGCGCGCGCGCGCGCGCG 160
QY 399 CCACCAAG 406
|||||
Db 159 CCAGGAAG 152

```

|          |                                   |
|----------|-----------------------------------|
| JOURNAL  | Pseudomonas aeruginosa library    |
| COMMENT  | J. Bacteriol. (2002) In press     |
|          | Contact: Chris K. Raymond         |
|          | Genome Center                     |
|          | University of Washington          |
|          | Box 352145, Seattle, WA 98105-21  |
|          | Tel: 2062216954                   |
|          | Fax: 2066857244                   |
|          | Email: chraymond@u.washington.edu |
| FEATURES | Class: shotgun.                   |
|          | Location/Qualifiers               |
| source   | 1. .990                           |

FEATURES  
SOURCE

Class: Shotgun.  
Location/Qualifiers  
1. .990

```

Cross: shotgun:
Location/Qualifiers
  1..990
    /organism="Pseudomonas aeruginosa"
    /mol_type="genomic DNA"
    /strain="MSH"
    /db_xref="taxon:287"
    /clone="mah2_3292"
    /clone_lib="mah"
    /note="Environmental isolate. Whole genomic shotgun
    library."

```

## ORIGIN

| Query Match           | 6.2%           | Score 57.4   | DB 8     | Length 990 |
|-----------------------|----------------|--|----------|------------|
| Best Local Similarity | 53.3%          | Pred. No. 4.6e-05  |          |            |
| Matches 121           | Conservative 0 | Mismatches 106   | Indels 0 | Gaps 0     |
| QY                    | 35             | TCGACGACCTACGCAGCTTCATTTTCAGTCGGCTCAATCAGGCCACTCACCAGAAATCGCGC | 94       |            |
| DB                    | 258            | TCGACCTGCTCCGGGTATTCGAGTTCGCGCGCCGCCAACTCAGTTTCACCGCTCGCGCGC   | 317      |            |
| QY                    | 95             | AAAGATTAGGATCCCGAGCCCAACATTTCCAGACGAATCAGCGGAGTGGAAAAACACG     | 154      |            |
| DB                    | 318            | AGGAACTGGGGACCAACCAGCCTCGGCTCAGCCAGCAGATCAAAACGCCCTGGAGAAAGAGC | 377      |            |
| QY                    | 155            | CAGSCACCCCACTTTTCAGCGCGCGCGCCGCAACTCGTCCTCAACCAACGAGGCCACG     | 214      |            |
| DB                    | 378            | TGGGACCCCTCTGTTTCAGCGCGTTTCATCGAGGCATCGTGCTCACCAGCCCGCGCCAGC   | 437      |            |
| QY                    | 215            | CTTCTCTCAACCAACCGCAGCGCCATCGTCGAGAAATTCACATCCGCGC              | 261      |            |
| DB                    | 438            | TTCTGCTCCGCGACGTGGCGCGCGGCTTCGAGAGCATCGATGCCGCGC               | 484      |            |

RESULT 7  
 BZ563333  
 LOCUS  
 DEFINITION  
 1077 bp DNA linear GSS 17-DEC-2002  
 pac82-164\_4204.v2 pac82-164 *Pseudomonas aeruginosa* genomic clone  
 pac82-164\_4204, genomic survey sequence.  
 ACCESSION  
 BZ563333  
 VERSION  
 BZ563333.1 GI:27186978  
 KEYWORDS  
 GSS.  
 SOURCE  
*Pseudomonas aeruginosa*  
 ORGANISM  
*Pseudomonas aeruginosa*  
 Bacteria; Proteobacteria; Gammaproteobacteria; *Pseudomonadales*;  
*Pseudomonadaceae*; *Pseudomonas*.  
 REFERENCE  
 1 (bases 1 to 1077)  
 Spencer D.H., Raymond C.K., Smith E.E., Sims E.E., Hastings M.,  
 Burns J.L., Kaul R. and Olsen M.V.  
 TITLE  
 Whole-Genome-Sequence variation among multiple isolates of  
*Pseudomonas aeruginosa* library  
 J. Bacteriol. (2002) In press  
 COMMENT  
 Contact: Chris K. Raymond  
 Genome Center  
 University of Washington  
 Box 352445, Seattle, WA 98105-2145, USA  
 Tel: 2062221854  
 Fax: 2066857244  
 Email: craymond@u.washington.edu  
 Class: shotgun.

FEATURES  
source

```

Cross. smogun:
  Location/Qualifiers
    1..1077
      /organism="Pseudomonas aeruginosa"
      /moltype="genomic DNA"

```

```
/strain="2-164"  
/db_xref="taxon:287"  
/clone="pac82-164_4204"  
/clone_lib="pac82-164"  
/note="clinical isolate 2-164 Whole genomic shotgun  
library."
```

## ORIGIN

| Query Match           | 6.1%           | Score 56.8  | DB 8     | Length 1077 |
|-----------------------|----------------|---|----------|-------------|
| Best Local Similarity | 46.0%          | Pred. No. 7e-05   |          |             |
| Matches 190           | Conservative 0 | Mismatches 223  | Indels 0 | Gaps 0      |
| Qy                    | 39             | CGACCTACGAGCTTCATTTTCAGTCGCTCAATCAGGCCACTCACCAGAACTGCCGAAG      | 98       |             |
| Db                    | 144            | CCAACCTCGGTATTTCTGTCGGCTCTACGAGGAAGCCACGTCGGCGCGCGCCGGAACG      | 203      |             |
| Qy                    | 99             | ATTAGGCATCCCGACGCCACACTTTTCCAGACGAATCAGCCGAGTGGAAAAACACGCAGG    | 158      |             |
| Db                    | 204            | CCTGGCCCTGTGCAACCGGCCCTCTCGCAGCAGATCCGCAGTTGGAGCAGAACTCTCGA     | 263      |             |
| Qy                    | 159            | CACCCCACTTTTCGACGCGCGCGGCCCAACTCGTCTCAACCAACGAGGCGCAGCGCTT      | 218      |             |
| Db                    | 264            | CGTCTCGCTGTTCCAGGCGCAGCGCAGCGCTGCTGCCGACCGTCGCGCGCCACACCGT      | 323      |             |
| Qy                    | 219            | CCTCAACACCGCCAGCGCCACTCGTCGCAGAAATTCAACTCGCGCCCAACTGAAATCAAAACG | 278      |             |
| Db                    | 324            | GTACAAACCATGCGGTACCGCTGCTGGAGGGATTGGAAAGGCCCGCAGGCGCATCGCGGT    | 383      |             |
| Qy                    | 279            | CCTCATGGACCCAGAAAAAGGCAATCCGACTGGAATTCTATGATTTCTTTGGGCACTTG     | 338      |             |
| Db                    | 384            | ATTCCGCGGGCAGACCCCGCGCAGCCTTGGGGATCGGGGTACTGCAGACGGTCAACGCCAG   | 443      |             |
| Qy                    | 339            | GATTGGTCCCGAACTTATTCGGAACATTTCCGCGCCGGAACACCCCAACGTAGAATTTCCAAC | 398      |             |
| Db                    | 444            | CCTGGTCCGTACATGTTGAAACGCTCTGCAGCCCGTGCAGCCGATWTGCGGGTGCAGAT     | 503      |             |
| Qy                    | 399            | CCACCAAGCGCAGCAATGCTCCTGGTAGATCGTGTGTTTGGCTGATGAAACTG           | 451      |             |
| Db                    | 504            | CTACGAACTCTCGGAAATGACATTCGAGCGCGCGCTGCTCACCGGAAACCTG            | 556      |             |

## RESULT 8

```

CN822018      717 bp      mRNA      linear      EST 02-JUN-2004
Oa splbn_10N24 M13reverse Sheep spleen\brain pSport1 library Ovis
aries cDNA clone Oa_splbn_10N24 5', mRNA sequence.

CN822018      GI:47950087
EST.
Ovis aries (sheep)
Ovis aries
Ovis aries
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Caprinae; Ovis.

REFERENCE
1 (bases 1 to 717)
Gossner,A. and Hopkins,J.
Ovine spleen\brain cdna library
Unpublished (2004)

CONTACT: J Hopkins
Veterinary Biomedical Sciences
University of Edinburgh
Summerhall Square, Edinburgh, EH9 1QH.
Email: j.hopkins@ed.ac.uk
Plate: 10 row: N column: 24
Seq primer: M13reverse
High quality sequence start: 6
High quality sequence stop: 550.
Location/Qualifiers
1..717
/organism="Ovis aries"
/mol_type="mRNA"
/db_xref="taxon:9940"
/clone="Oa splbn_10N24"

FEATURES
source
```

|                       |                 |   |           |             |
|-----------------------|-----------------|---|-----------|-------------|
| Query Match           | 6.0%;           | Score 56.2;   | DB 7;     | Length 768; |
| Best Local Similarity | 48.9%;          | Pred. No. 9.7e-05;  |           |             |
| Matches 151;          | Conservative 0; | Mismatches 158;   | Indels 0; | Gaps 0;     |
| QY                    | 29              | TGGGAATCGACGACCTACGGCAGCTTCATTTCAGTCGCTCAATCATAGGCCACCTCACCGAAA | 88        |             |

Db 438 ACACCTCGGTACGCGCTGTTTATCCGCGCAGCGAGTACGCTGACCGAGCGG 497  
 Qy 209 GCCAGCGCTTCTCAACACGCGCGCATGTCGCGAGATTCAATTCGCGCGCAATG 268  
 Db 498 GCGCGCGATTGTTATCGAGGCGAGGCGACCGTGCAGTTCAGTTTCCCGCAGCA 557  
 Qy 269 AAATCAACGCTCATGACCGCCAGAAAAGGCACATCCGACTGCACTTCATGATTCCT 328  
 Db 558 ATGCCAGCGCGCGCGGTGAGATTGGTTCACATTGAATTGGGTACGTTGCGTCGG 617  
 Qy 329 TGGGCACTT 337  
 Db 618 CGGTCTATT 626

RESULT 11  
 BZ424306/c  
 LOCUS BZ424306 550 bp DNA linear GSS 10-DEC-2002  
 DEFINITION id59d05.g1 WGS-SbicolorF (DH5a methyl filtered) Sorghum bicolor  
 genomic clone id59d05 5', genomic survey sequence.

ACCESSION BZ424306  
 VERSION BZ424306.1 GI:26374408  
 KEYWORDS GSS.

SOURCE Sorghum bicolor (sorghum)

ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 550)  
 AUTHORS Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,  
 Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,  
 Zucavern,T., Palmer,I., McCombie,W.R. and Martienssen,R.A.  
 Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)  
 Unpublished (2002)

TITLE Lita Annenberg Hazen Genome Sequencing Center  
 JOURNAL Cold Spring Harbor Laboratory  
 COMMENT PO Box 100, Cold Spring Harbor, NY 11724, USA  
 Tel: 516 367 8884  
 Fax: 516 367 8874

FEATURES  
 source  
 1..550  
 Location/Qualifiers  
 /organism="Sorghum bicolor"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4558"  
 /clone="id59d05"  
 /lab\_host="DH5a"  
 /clone\_lib="WGS-SbicolorF (DH5a methyl filtered)"  
 /note="Site 1: Xba I; Site 2: Xba I; The vector was  
 digested with XbaI and one nucleotide was added by fill in  
 in the recessive 3' end. The genomic DNA was nebulized,  
 end repaired, adaptor ligated and size fractionated using  
 sephadex. The resulting fragments were between 0.8 and 3  
 kb and were cloned into the vector (.x/y reads in M13mp19,  
 .b/g reads in pUC19). The same ligation was transformed  
 into DH5a."

## ORIGIN

Query Match 6.0%; Score 55.8; DB 8; Length 550;  
 Best Local Similarity 52.9%; Pred. No. 0.00012;  
 Matches 120; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy 16 GACGCGGAGACCTCGGATCGACGACCTAGCAGCTTCATTTCAGTCGCTCAATCAGGC 75  
 Db 501 GACGTTTGTATATGAAACTTCGCACCTCGGTTATTTCGTTCGCGTTGCTGAAACGAG 442  
 Qy 76 CACCTCACCGAAACTGCCGAAAGATTAGGCATCCGCGAGCCACACTTTCAGAGCAATC 135

Db 441 CACTTACCCGCGCGCGAGTTCTGCGCATGTCGAGCCACCGCTCAGCGCAATC 382  
 Qy 136 AGCCAGTGTGAAAAACAGCAGGACCCCACTTTTCGACCGCGCGCCGCAAACTCGTC 195  
 Db 381 CAGCGGCTGGAGACGAGATCGGCACGCGCTGTTGTAAGCGCCTGACGCGCGCGTGAA 322  
 Qy 196 CTCAACCAACAGCGCCAGCCCTTCTCAACACGCGCAGCGCATCGT 242  
 Db 321 CTGACCGATGCGGCGCGCGCTTCTGCCGATGCAATAGCGATTGT 275

RESULT 12  
 BZ566287

LOCUS BZ566287 817 bp DNA linear GSS 17-DEC-2002  
 DEFINITION pacs2-164\_6003.y2 pacs2-164 Pseudomonas aeruginosa genomic clone  
 pacs2-164\_6003, genomic survey sequence.

ACCESSION BZ566287  
 VERSION BZ566287.1 GI:27195555

KEYWORDS GSS.  
 SOURCE Pseudomonas aeruginosa

ORGANISM Pseudomonas aeruginosa  
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 817)  
 AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,  
 Burns,J.L., Kaul,R. and Olsen,M.V.

TITLE Whole-Genome-Sequence Variation among multiple isolates of  
 Pseudomonas aeruginosa library

JOURNAL J. Bacteriol. (2002) In press

COMMENT Contact: Chris K. Raymond

Genome Center  
 University of Washington  
 Box 352145, Seattle, WA 98105-2145, USA  
 Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun

FEATURES  
 source  
 1..817  
 Location/Qualifiers

/organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="2-164"  
 /db\_xref="taxon:287"  
 /clone="pacs2-164\_6003"  
 /clone\_lib="pacs2-164"  
 /note="clinical isolate 2-164 Whole genomic shotgun  
 library."

## ORIGIN

Query Match 6.0%; Score 55.6; DB 8; Length 817;  
 Best Local Similarity 50.6%; Pred. No. 0.00015;  
 Matches 161; Conservative 0; Mismatches 154; Indels 3; Gaps 1;

Qy 89 CTGCGGAAAGATTAGGCATCCCGCAGCCACACTTTCAGACGAATCAGCGAGTGGAAA 148  
 Db 125 CCGTGTAGCGCTGCATGCGCCAGCACCCTGACCATCGCATCGCGCGGTGGAG 184  
 Qy 149 AACACGAGGACCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACGAG 208  
 Db 185 AAGAAGTGGCGGCAGCTGCTGGAGCGTGACACCGTATCCACCGCTCAGCGCGCAG 244  
 Qy 209 GCCAGCGCTTCTCAACACGCGCAGCGCATCGTGGAGATTCACTCCGCGCAATG 268  
 Db 245 GCGCGCGCTTCTCGAGAAAGCCCGCGCACCTTGGGCCAGCGCAACCGCTGGCGG 304  
 Qy 269 AAATCAACGCTCATGACCCAGAAAAAGGCACAAATCCGACTGCACTTCATGATTCCT 328  
 Db 305 ---CCGCGGAGCGCGCCAGCGAGCGAGAAAGTTGGCGCTGCGCTGCTCGACGCA 361  
 Qy 329 TGGGCACTTGATGGTCCCGAACTTATCCGAACTTATCCGCGCGGAAACACCCCAACGTA 388  
 Db 362 CCATCAATATTCTGCTACCGCGCATACTCCAGGCGTTCCGCAACAGCATGGCGCGCTGG 421

RESULT 14  
BZ579898

TITLE  
JOURNAL  
COMMENT

Zutavern, T., Palmer, L., McCombie, W.R. and Martienssen, R.A.  
Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)  
Unpublished (2002)  
Contact: W. Richard McCombie  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org  
Plate: id59 row: d column: 05  
Seq primer: -21M13UnivFwd  
Class: shotgun  
High quality sequence stop: 451.

Location/Qualifiers  
1. 451  
/organism="Sorghum bicolor"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4558"  
/clone="id59d05"  
/lab\_host="DH5a"  
/clone\_lib="WGS-SbicolorF (DH5a methyl filtered)"  
/note="Site 1: Xba I; Site 2: Xba I; The vector was  
digested with XbaI and one nucleotide was added by fill in  
in the recessive 3' end. The genomic DNA was nebulized,  
end repaired, adaptor ligated and size fractionated using  
sephadex. The resulting fragments were between 0.8 and 3  
kb and were cloned into the vector (.x/y reads in M13mp19,  
.b/g reads in pUC19). The same ligation was transformed  
into DH5a."

FEATURES  
source

ORIGIN

Query Match 5.9%; Score 54.8; DB 8; Length 451;  
Best Local Similarity 53.2%; Pred. No. 0.00022;  
Matches 116; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 25 GACCTGCGAATCGACGACCTACGACGCTTTCATTTTCAGTCGCTCAATCAGGCGACCTCACC 84  
|||  
Db 9 GATATGGAACCTCGCCACCTCGGTATTTCGTTGCGCTTGCTGAACGCGACGACTTTACC 68  
|||  
QY 85 GAACTGCGGAAGATTAGGATCCCGCAGCCCACTTTCCAGACGAATCAGCGAGTG 144  
|||  
Db 69 CGCGCGCGGAGTTGCTGGGATGTCGAGCCACCGCTCAGCCAGCAATCCAGCGGCTG 128  
|||  
QY 145 GAAAAACACGAGGACGCCCTTTTCGACCGCGCGCGCAACTCGTCTCAACCAA 204  
|||  
Db 129 GAGCAGGAGATCGGCACCGCTGTTGTTGAAGCGCTGACGCGCGGTGGAACGACCGAT 188  
|||  
QY 205 CGAGGCCACGCTTCTCAACCGCCAGCGCCATCGT 242  
|||  
Db 189 GCGGCGCGGCTTCTGCGCGATGCAATAGGATTGT 226  
|||

Search completed: April 13, 2005, 01:59:27  
Job time : 3109 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2005, 23:32:55 ; Search time 206 Seconds  
(without alignments)

7387.075 Million cell updates/sec

Title: US-09-826-909-1\_COPY\_232\_1161

Perfect score: 930

Sequence: 1 gtgaccatgggcaacgacgg.....ggatgcattagaagagggc 930

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*\*

- 1: /cgn2.6/prodata/1/ina/5A COMB.seq:\*\*
- 2: /cgn2.6/prodata/1/ina/5B COMB.seq:\*\*
- 3: /cgn2.6/prodata/1/ina/6A COMB.seq:\*\*
- 4: /cgn2.6/prodata/1/ina/6B COMB.seq:\*\*
- 5: /cgn2.6/prodata/1/ina/PCTUS COMB.seq:\*\*
- 6: /cgn2.6/prodata/1/ina/backfiles1.seq:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| C 1        | 129.6 | 13.9        | 77536  | 4  | US-09-410-551B-1     |
| C 2        | 129.6 | 13.9        | 77536  | 4  | US-09-940-316B-1     |
| C 3        | 68    | 7.3         | 927    | 4  | US-09-252-991A-12324 |
| C 4        | 68    | 7.3         | 972    | 4  | US-09-252-991A-12096 |
| C 5        | 68    | 7.3         | 1482   | 4  | US-09-252-991A-12421 |
| C 6        | 61.4  | 6.6         | 987    | 4  | US-09-252-991A-7082  |
| C 7        | 61.4  | 6.6         | 1485   | 4  | US-09-252-991A-7058  |
| C 8        | 61.4  | 6.6         | 1518   | 4  | US-09-252-991A-7044  |
| C 9        | 59.2  | 6.4         | 1035   | 4  | US-09-252-991A-13904 |
| C 10       | 57.8  | 6.2         | 657    | 4  | US-09-252-991A-3008  |
| C 11       | 57.8  | 6.2         | 1074   | 4  | US-09-252-991A-2667  |
| C 12       | 57.4  | 6.2         | 471    | 4  | US-09-252-991A-5341  |
| C 13       | 57.4  | 6.2         | 552    | 4  | US-09-252-991A-5279  |
| C 14       | 57.4  | 6.2         | 909    | 4  | US-09-252-991A-5304  |
| C 15       | 57.4  | 6.2         | 2592   | 4  | US-09-252-991A-5260  |
| C 16       | 56.4  | 6.1         | 771    | 4  | US-09-252-991A-3268  |
| C 17       | 56.4  | 6.1         | 888    | 4  | US-09-252-991A-3414  |
| C 18       | 56.4  | 6.1         | 1197   | 4  | US-09-252-991A-3336  |
| C 19       | 55    | 5.9         | 7295   | 4  | US-09-902-540-924    |
| C 20       | 54.8  | 5.9         | 711    | 4  | US-09-252-991A-12330 |
| C 21       | 54.8  | 5.9         | 930    | 4  | US-09-252-991A-12091 |
| C 22       | 54.6  | 5.9         | 19394  | 4  | US-09-902-540-1172   |
| C 23       | 54.2  | 5.8         | 945    | 4  | US-09-252-991A-2028  |
| C 24       | 54.2  | 5.8         | 984    | 4  | US-09-252-991A-3477  |
| C 25       | 54.2  | 5.8         | 1392   | 4  | US-09-252-991A-3486  |
| C 26       | 54    | 5.8         | 921    | 4  | US-09-252-991A-3438  |
| C 27       | 54    | 5.8         | 1263   | 4  | US-09-252-991A-3462  |

|      |      |     |       |   |                      |                   |
|------|------|-----|-------|---|----------------------|-------------------|
| 28   | 53.8 | 5.8 | 885   | 4 | US-09-902-540-8724   | Sequence 8724, Ap |
| 29   | 53.6 | 5.8 | 891   | 4 | US-09-902-540-3768   | Sequence 3768, Ap |
| 30   | 53.6 | 5.8 | 936   | 4 | US-09-902-540-3745   | Sequence 3745, Ap |
| 31   | 53.6 | 5.8 | 17897 | 4 | US-09-902-540-1182   | Sequence 1182, Ap |
| C 32 | 53.6 | 5.8 | 20303 | 4 | US-09-902-540-1183   | Sequence 1183, Ap |
| 33   | 53.4 | 5.7 | 936   | 4 | US-09-489-039A-2270  | Sequence 2270, Ap |
| C 34 | 53.2 | 5.7 | 732   | 4 | US-09-252-991A-6298  | Sequence 6298, Ap |
| C 35 | 53.2 | 5.7 | 855   | 4 | US-09-252-991A-11951 | Sequence 11951, A |
| C 36 | 53.2 | 5.7 | 897   | 4 | US-09-489-039A-1429  | Sequence 1429, Ap |
| C 37 | 53.2 | 5.7 | 945   | 4 | US-09-252-991A-6167  | Sequence 6167, Ap |
| C 38 | 53.2 | 5.7 | 978   | 4 | US-09-252-991A-12061 | Sequence 12061, A |
| 39   | 53.2 | 5.7 | 1602  | 4 | US-09-252-991A-6382  | Sequence 6382, Ap |
| 40   | 53   | 5.7 | 621   | 4 | US-09-252-991A-1463  | Sequence 1463, Ap |
| C 41 | 53   | 5.7 | 939   | 4 | US-09-252-991A-1511  | Sequence 1511, Ap |
| C 42 | 53   | 5.7 | 3429  | 4 | US-09-252-991A-1663  | Sequence 1663, Ap |
| C 43 | 52.6 | 5.7 | 936   | 4 | US-09-252-991A-12221 | Sequence 12221, A |
| C 44 | 52.4 | 5.6 | 948   | 4 | US-09-252-991A-12230 | Sequence 12230, A |
| C 45 | 52   | 5.6 | 291   | 4 | US-09-252-991A-6225  | Sequence 6225, Ap |

ALIGNMENTS

RESULT 1

US-09-410-551B-1/c  
; Sequence 1, Application US/09410551B  
; Patent No. 6503737  
; GENERAL INFORMATION:  
; APPLICANT: KOSAN BIOSCIENCES, Inc.  
; APPLICANT: REEVES, CHRISTOPHER  
; APPLICANT: CHU, DANIEL  
; APPLICANT: KHOSLA, CHAITAN  
; APPLICANT: SAMI, DANIEL  
; APPLICANT: WU, KAI  
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA  
; FILE REFERENCE: 30062-20026.00  
; CURRENT APPLICATION NUMBER: US/09/410,551B  
; CURRENT FILING DATE: 1999-10-01  
; PRIOR APPLICATION NUMBER: US 60/139,650  
; PRIOR FILING DATE: 1999-06-17  
; PRIOR APPLICATION NUMBER: US 60/123,810  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 60/102,748  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 77536  
; TYPE: DNA  
; ORGANISM: Streptomyces hygroscopicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (52275)...(71465)  
US-09-410-551B-1

Query Match 13.9%; Score 129.6; DB 4; Length 77536;  
Best Local Similarity 50.7%; Pred. No. 4.3e-28;  
Matches 374; Conservative 0; Mismatches 349; Indels 15; Gaps 2;

|    |      |  |      |
|----|------|--|------|
| QY | 27   | CCTGCGAATCGACGACCTACGAGCTTCATTTTCAGTCGCTCAATCAGGCCACCTCACCGA | 86   |
| Db | 5489 | CATGAGATACGGAGTTGCACTGTTTACGGCCCTTGGCCGAGCAGGACACGTGACGAT    | 5430 |
| QY | 87   | AACTCCGGAAGATTAGGCATCCCGAGCCCACTTTTCCAGCAGNATCAGCGAGTGA      | 146  |
| Db | 5429 | CACCGCGAAGCGCTCAACATCTCCAGCCGACCTGTCTCAGGCGCATCCGCGGTGA      | 5370 |
| QY | 147  | AAACACGCGAGCACCCCACTTTTTCGACGCGCGCGCAAACTGCTCTCAACCAACG      | 206  |
| Db | 5369 | GGCAGAGTGGCGGTGCGCTGTTTCACCGGACACAGAACCGGCTCCGGCTCAACAGTA    | 5310 |
| QY | 207  | AGGCCACGCGCTTCTCTCAACACGCGCGCCATCTGTCGAGAAATTCACCTCCGCGCAAC  | 266  |

|      |  |      |
|------|--|------|
| 5309 | CGGCGAGGCTTTCGCGCCACGCGCCCTGCGCGGATCAGTGAGTCTTCGCGCGCGGACGA    | 5250 |
| 267  | TGAATCAAAAGCCGCTCATGACCCAGAAAAAGGCACAATCCGACTGACATTCATGCATTC   | 326  |
| 5249 | GCGCATCACCGCCCTGATCGACCCGACACGCGACCGTGGCACTGGTTTCTGCACTC       | 5190 |
| 327  | CTTGGGCACTTGGATGCTCCCGAACTTATCCGAAACATTCGCGCGCGAAACACCCCAACGT  | 386  |
| 5189 | CTACGGACCTGGCTGGTGGCCGACCTCTCGCGGGTACCGCGGGTGGCTCCCGGCAC       | 5130 |
| 387  | AGAAATCCAACTCCACCAAGCGGACGAATGCTCTGATAGATCGTGTTCGCTGATGA       | 446  |
| 5129 | CACCTTCGAATCTGCGCGGACGCGCGCGACGCCGTGGTGCAGACCTCCGCAACAGCAG     | 5070 |
| 447  | AACTGACCTCGCATTAGTTTGGCCCCAAACCTCGCGAGGTTTGTAACCTCTTTAGGTGGGCG | 506  |
| 5069 | GCTGGACCTGGGGCTGACCAAGCCACGACCGCGCGGG-----ACGATGTGGAGTGGAC     | 5016 |
| 507  | GCCACTGCTTCGTCAAGACTTGGCCCTAGCTGTTCGCGAGATCACCGGCTTGCTTCCTT    | 566  |
| 5015 | ACCACTCGGAGCGAGGCACTGTGCTGTCTGCTCCACCCGCGACACCGGTGGCC-----     | 4961 |
| 567  | TTCTGGCCAAAGGAGAAATGCGCTTGATTACTGCGGCGGAAAGAACCTTTCGTGCGATGCG  | 626  |
| 4960 | ----AGACGCGCGCGGTGCGGACCGCGGAGCTGGCCGATGAGCGCTTCGTGSCACTGGA    | 4905 |
| 627  | AGCAGGTTTCGGCACCCGACTCTCTCATGATGATATAGCCGAGAGACCGGTTTGTGTTC    | 686  |
| 4904 | GCGCGTCTTCGGACTGCGCGCAGATCACCGACCGGCTGTGCGCGCGCGCCGGAATTCGTGCC | 4845 |
| 687  | CAATGTGGTTTTCGAATCCATGAACTCACACCGTCGACGGGCTGTGACGCGCAGCTCT     | 746  |
| 4844 | GCGGTCGTCTGTGAGAGACCCGAGCTTGAGCAGCTCGGGGCCCTGTGTCGCCCGCGCCT    | 4785 |
| 747  | CGCGGTTGGTGTGGTTCC   | 764  |
| 4784 | CGGTGTGGCGTGCCTCC  | 4767 |

RESISTANCE

```

RESUL1 2
US-091-940-316B-1/c
; Sequence 1, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDES ENCODING T
; TITLE OF INVENTION: GENE CLUSTER
; FILE REFERENCE: 30062-20026.11
; CURRENT APPLICATION NUMBER: US/09/940,316B
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/410,551
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 77536
; TYPE: DNA
; ORGANISM: Streptomyces hygroscopicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52275)...(71465)

```



; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 12324  
; LENGTH: 927  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12324

Query Match 7.3%; Score 68; DB 4; Length 927;  
Best Local Similarity 49.2%; Pred. No. 2.4e-10;  
Matches 179; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 38 AGCACTACGAGCTTCATTTCAGTCGCTCAATCAGGCCACCTCACCAGAACTGCCGAAA 97  
DB 44 ACAGCCTCAACACCTTCATCGCATTGCGAGACCGGACGCTTCCTCGAAGCGCGGAGC 103  
QY 98 GATTAGGCATCCGAGCCACACATTCAGAGCATCAGCGAGTGGGAAACACGCGAG 157  
DB 104 GCTGACCTGACCCCAACCGCGGTGAGCAAGCGCATCGCCCTCGAAGACGAGCTCA 163  
QY 158 GCACCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACGAGGCCACGCT 217  
DB 164 ATGCGCGGTGTTTCGACCGGTGCGCGGAGGTCAACTCAGGAGCGCGCGCGCC 223  
QY 218 TCTCAACCAACGAGCGGATTCGTCGAGAAATCAACTCCCGCAACTGAAATCAAAAC 277  
DB 224 TGTGCGCGCGCTACCACTGCTCAAGTGTCTGACGACACCCCGCGCGGCTGAACA 283  
QY 278 GCTCATGACCCAGAAAGGACACATCCGACTGACATTCATGATTCCTTTGGGACTT 337  
DB 284 ACCTCACCGCGAGGTGACGCGCGGCTGCTCTGCGCCACGACCAACCATCGGCTGC 343  
QY 338 GGATGGTCCCGCAACTTATCCGAAACATTCGCGCGGCAACACCCCAAGTAGAATCCAAC 397  
DB 344 ATGCGCTGCGCGGTGCTGCGGCTTACCGGCGCATCCGAGGTGGCGCTGGATA 403  
QY 398 TCCA 401  
DB 404 TCCA 407

RESULT 4  
US-09-252-991A-12096/c  
; Sequence 12096, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 12096  
; LENGTH: 972  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12096

Query Match 7.3%; Score 68; DB 4; Length 972;  
Best Local Similarity 49.2%; Pred. No. 2.4e-10;  
Matches 179; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 38 AGCACTACGAGCTTCATTTCAGTCGCTCAATCAGGCCACCTCACCAGAACTGCCGAAA 97  
DB 890 ACAGCTCAACACCTTCATCGCATTGCGAGACCGGACGCTTCTCCGAGCGCGGAGC 831  
QY 98 GATTAGGCATCCGAGCCACACATTCAGAGCAATCAGCGAGTGGGAAACACGCGAG 157

DB 830 GCTGCACTGACCCCAACCGCGGTGAGCAAGCGCATCGCCGCTCGAACAACGAGCTCA 771  
QY 158 GCACCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACGAGGCCGAGCTT 217  
DB 770 ATGCCCCGCTGTTTCGACCGGTGCGCGGAGGTCAACTCACCAGGCCGCGCGGCC 711  
QY 218 TCTCAACCAACGCGGAGCGCATTCGTCGAGAAATTCAACTCCCGCGCAACTGAAATCAAAAC 277  
DB 710 TGTGCGCGCGCTTACAGTTGCTCAAGTGTCTGAGACACCGCGCGGCGCTGAACA 651  
QY 278 GCTCATGACCCAGAAAGGACACATCCGACTGACATTCATGATTCCTTTGGGACTT 337  
DB 650 ACCTCACCGCGAGGTGACGCGCGGCTGGTCTTGCGCCACCAAGCCACCATCGGCTGC 591  
QY 338 GGATGGTCCCGCAACTTATCCGAAACATTCGCGCGGCAACACCCCAAGTAGAATCCAAC 397  
DB 590 ATGCGCTGCGCGGTGCTGCGGCTTCAACCGCGCATCCGAGGTGGCGCTGGATA 531  
QY 398 TCCA 401  
DB 530 TCCA 527

RESULT 5  
US-09-252-991A-12421  
; Sequence 12421, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 12421  
; LENGTH: 1482  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12421

Query Match 7.3%; Score 68; DB 4; Length 1482;  
Best Local Similarity 49.2%; Pred. No. 3e-10;  
Matches 179; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 38 ACAGCTACGAGCTTCATTTCAGTCGCTCAATCAGGCCACCTCACCAGAACTGCCGAAA 97  
DB 99 ACAGCTCAACACCTTCATTCGCGATTCGCGAGACCGGAGCTTCTCCGAGCGCGGAGC 158  
QY 98 GATTAGGCATCCGAGCCACACATTCAGAGCAATCAGCGAGTGGGAAACACGCGAG 157  
DB 159 GCTGCACTGACCCCAACCGCGGTGAGCAAGCGCATCGCCGCTCGAAGACGAGCTCA 218  
QY 158 GCACCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACGAGGCCAGCTT 217  
DB 219 ATGCGCGGTGTTTCGACCGGTGCGCGGAGGTCAACTCACCAGGCGCGCGGCC 278  
QY 218 TCTCAACCAACGCGGCTTCGTCGAGAAATTCAACTCCGCGCAACTGAAATCAAAAC 277  
DB 279 TGTGCGCGGCTTACAGTTGCTCAAGTGTCTGAGACACCGCGCGGCGCTGAACA 338  
QY 278 GCTCATGACCCAGAAAGGACCAATCCGACTGAGCTTTCATGATTCCTTTGGGACTT 337  
DB 339 ACCTCACCGGAGGTGACGCGCGGCTGCTGCGCCACCAAGCCACCATCGGCTGC 398  
QY 338 GGATGGTCCCGAACTTATCCGAAACATTCGCGCGGCAACACCCCAAGTAGAATCCAAC 397  
DB 399 ATGCGCTGCGCGGTGCTGCGGCTTCAACCGGCGCATCCGAGGTGGCGCTGGATA 458

QY 398 TCCA 401  
|||  
Db 459 TCCA 462

## RESULT 6

US-09-252-991A-7082  
; Sequence 7082, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7082  
; LENGTH: 987  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7082

Query Match 6.6%; Score 61.4; DB 4; Length 987;  
Best Local Similarity 47.7%; Pred. No. 2.5e-08;  
Matches 179; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 27 CTGCGAATCGACGACCTACGAGCTTCATTTTCAGTCGCTCAATCAGCGCACTCACC GA 86  
|||  
Db 78 CTGCGTTTCGACGATATCGAAGCTTCTCAAGCTCGCGCACTCAACAGCGTGACGTC 137  
|||  
QY 87 AACTCCGAAAGATTAGGCATCCCGAGCCCACTTTTCCAGACGAATCAGCCGAGTGGA 146  
|||  
Db 138 GCGCGCGAGCGCATGAGCTGTCCAAGTCGGTGATCAGCAAGCGTGTGACGATCTCGA 197  
|||  
QY 147 AAAACACGAGCAGCCCACTTTTCGACGCGCGCGCAACTCGTCTCAACCAAG 206  
|||  
Db 198 ACGAGCGCTCGCGCTCGCGCTCTATGCGACCAAGCGCAAGCTGAGCCGACCGAGGC 257  
|||  
QY 207 AGGCCAGCGCTTCCTCAACAGCGCAGCGCCATCGTCGAGAAATCAACTCGCGCAAC 266  
|||  
Db 258 CGGGCATTTCTTACAGGCGCGCAAGGCTCGCTGCGAGAACTCAACAGCGCGCGGA 317  
|||  
QY 267 TGAATCAAAACGCTCATGGACCCAGAAAGGCAATCCGACTGGACTTTCATGCAATTC 326  
|||  
Db 318 GATTTCGCGCTCGCGAGAACGACCTGTGCGGGAATTTGGCATCGTCACCCCGATGAG 377  
|||  
QY 327 CTTGGCACTTGGATGGTCCCGAACTTATCCGAACATTCGCGCGGCAACACCCCAAGT 386  
|||  
Db 378 CTTGCGACCTCTCGCTGTGCGCGATCATCGCGCACTTCATGAGCCGACCCCGCGCT 437  
|||  
QY 387 AGAATTCCAACTCCA 401  
|||  
Db 438 GGAGATCGTCTGCA 452  
|||

## RESULT 7

US-09-252-991A-7058  
; Sequence 7058, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7058  
; LENGTH: 1485  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7058

Query Match 6.6%; Score 61.4; DB 4; Length 1485;  
Best Local Similarity 47.7%; Pred. No. 3e-08;  
Matches 179; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 27 CCTCGAATCGACGACCTACGAGCTTCATTTTCAGTCGCTCAATCAGCGCACTCACC GA 86  
|||  
Db 124 CTGCGTTTCGACGATATCGAAGCTTCTCTCAAGTCGCTGATCAGCAAGGCTGTGTCAGGGATCTCGA 183  
|||  
QY 87 AACTGCCGAAAGATTAGGCATCCCGAGCCCACTTTTCCAGACGAATCAGCCGAGTGGA 146  
|||  
Db 184 GCGCGCGAGCGCATGAGCTGTCCAAGTCGGTGATCAGCAAGGCTGTGTCAGGGATCTCGA 243  
|||  
QY 147 AAAACACGAGCAGCCCACTTTTTCGACGCGCGCGCAACTCGTCTCAACCAAG 206  
|||  
Db 244 ACGAGCGCTCGCGCTCGCGCTCTATGCGACCAAGCGCAAGCTGAGCGCGAGCGAGGC 303  
|||  
QY 207 AGGCCAGCGCTTCCTCAACAGCGCAGCGCCATCGTCGCAAGATTCACCTCGCGCAAC 266  
|||  
Db 304 CGGGCATTTCTTACAGGCGCGCAAGGCTCGCTGCGAGAACTCAACAAACGCGCGGA 363  
|||  
QY 267 TGAATCAAAACGCTCATGGACCCAGAAAGGCAATCCGACTGGACTTTCATGCAATTC 326  
|||  
Db 364 GATTTCGCGCTCGCGAGAACGACCTGTGCGGGAATTTGGCATCGTCACCCCGATGAG 423  
|||  
QY 327 CTTGGCACTTGGATGGTTCGCGAACTTATCCGAACATTCGCGCGGCAACACCCCAAGT 386  
|||  
Db 424 CTTGCGACCTCTCGCTGTGCGGATCATCGCGACTTCATGAGCCGACCCCGCGCT 483  
|||  
QY 387 AGAATTCCAACTCCA 401  
|||  
Db 484 GGAGATCGTCTGCA 498  
|||

## RESULT 8

US-09-252-991A-7044/c  
; Sequence 7044, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7044  
; LENGTH: 1518  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7044

Query Match 6.6%; Score 61.4; DB 4; Length 1518;  
Best Local Similarity 47.7%; Pred. No. 3.1e-08;  
Matches 179; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 27 CCTCGAATCGACGACCTACGAGCTTCATTTTCAGTCGCTCAATCAGCGCACTCACC GA 86  
|||  
Db 1429 CTGCGTTTCGACGATATCGAAGCTTCTCTCAAGTCGCGCAACTCAACAGCGTGACGTC 1370  
|||  
QY 87 AACTGCCGAAAGATTAGGCATCCCGAGCCCACTTTTCCAGACGAATTCAGCCGAGTGGA 146  
|||

Db 1369 GCGGCGCGGCGCATGAGCGCTGTCCAAGTCGCTGATCAGCAAGCGTGTGAGGATCTCGA 1310  
Qy 147 AAAACACGCGAGCAGCCCACTTTTCGACCGCGCGCGCAAACTCGTCTCAACCAACG 206  
Db 1309 AGACGCGCTCGGGTCCGCTGCTCTATCGACCGCGCAAGCTGAGCGGACCGAGGC 1250  
Qy 207 AGGCCACGCGCTTCTCAACACGCGCAGCGCCATCGTCGAGAAATCAACTCGCGCGCAAC 266  
Db 1249 CGGGCACTTTCTTACAAGCGCGCAAGCGCTCGCTGCGAGAACTCAACAACGCGCGGA 1190  
Qy 267 TGAATCAACGCTCATGAGCAGCAAGAAAGGCAACATCGACTGGACTTATGCAATTC 326  
Db 1189 GATTGTGCGCCTGCGCGAGAACGACCTGTGCGCGGAATTCGCGCATGTCACCCCGATGAG 1130  
Qy 327 CTGCGCACTTGGATGCTCCCGAACTTATCGAACATTCGCGCGGCAACACCCAACT 386  
Db 1129 CTTCGCGACCTTGTGCTGTGCGGATCATGCGCGACTTATGAGCGAGCAGCAGCGGCT 1070  
Qy 387 AGAATTCGAATCCA 401  
Db 1069 GGAGATCGTCTGCA 1055

## RESULT 9

US-09-252-991A-13904  
; Sequence 13904, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 13904  
; LENGTH: 1035  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-13904

Query Match 6.4%; Score 59.2; DB 4; Length 1035;  
Best Local Similarity 49.5%; Pred. No. 1.2e-07;  
Matches 182; Conservative 0; Mismatches 183; Indels 3; Gaps 1;

Qy 39 CGACCTACGCGCTTCATTTAGTCGCTCAATCAGCGCCACCTCACCGAACTGCCGAAAG 98  
Db 144 CAGTTCCGCGAGTTTCGTCGCGTTCCGAGAGCTGAGCTTCGCGCGCGCTGAGCG 203  
Qy 99 ATTAGCCTCCGCGACCCACACTTTCCAGACGAATCAGCGAGTGGAAAAACAGCGAG 158  
Db 204 CTGACATGCGCGCCAGCCAGCTGACCACTCGATCCGCGCGCTGGAAGAAAGTCGG 263  
Qy 159 CACCCCACTTTTCGACCGCGCGCGGCGAACTTCGCTCAACCAAGAGGCGACGCTT 218  
Db 264 CGCGAGCTTCTGGAGCGTGACAACTGATCCACCGCTGAGCGCGGCGGCGGCTT 323  
Qy 219 CTTCAACACGCGCGCGCATCGTCGAGAAATTCACCTCGCGCGCAACTGAAATCAACG 278  
Db 324 CTTCGAGAAAGCCGCGCGCACCTTCGCGCGCAAGCGACCTCGCGG---CGCGCG 380  
Qy 279 CTTCAAGACCCAGAAAAAGGCAATTCGACTGGAATTCATGCAATTCCTTGGCACTTG 338  
Db 381 CAGCGCGCGCGAGCGAGCAAGCTTTCGCTGCGGTTCTGCGACAGCACCATCATAT 440  
Qy 339 GATGTGCGCGCACTTATCCGAACATTCGCGCGCGCAACCCCAACGTAATTCACCT 398  
Db 441 TCTGCTACCGCGCATACTCCAGGCGTTCCGCCACAGCATGCGCGCTGGAATTCAGCT 500

Qy 399 CCACCAAG 406  
Db 501 CCAGGAAG 508

## RESULT 10

US-09-252-991A-3008/c  
; Sequence 3008, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 3008  
; LENGTH: 657  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-3008

Query Match 6.2%; Score 57.8; DB 4; Length 657;

Best Local Similarity 52.7%; Pred. No. 2.5e-07;  
Matches 125; Conservative 0; Mismatches 112; Indels 0; Gaps 0;  
Qy 27 CCGCGAATCGAGCAGCTACGAGCTTCATTTTCAGTCGCTCAATCAGCGCACCTCACCGA 86  
Db 595 CCGCGCTTGAAGACATACAGCGCTTCTCAAGCTGTCGAGCTGGGAGCATCAGCG 536  
Qy 87 AACTCCGAAAGATTAGGCATCCCGAGCCCACTTTCCAGACGAATCAGCGAGTGA 146  
Db 535 CGCGCGCGAAGCATGAGCTGTCCAAGTCCGATCAGCAACGCGTCAGCGACCTGA 476  
Qy 147 AAAACACGCGCAGCCCACTTTTCGACCGCGCGCGCAACTCTCTCAACCAACG 206  
Db 475 GCGCCACCTCGCGCGCTCTCTACCGCTCGACGCGCAACGTCGAGCGCGAGGC 416  
Qy 207 AGGCCACGCGCTTCTCAACACGCGCGCGCATCGTCGAGAAATTCACCTCGCGCG 263  
Db 415 CGCGCGCTTCTTCAAGTCGCGCAAGCGCTCGCTGAGGACCTGGAACAGCGCGC 359

## RESULT 11

US-09-252-991A-2667  
; Sequence 2667, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 2667  
; LENGTH: 1074  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1046)  
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.



QY 215 CCTTCTCAACACGCGCCATCGTCGAGAAATTCAACTCCGCC 261  
 Db 203 TTCTGCTCCGCGGCGGCTCGAGACGATCGCCGC 249

RESULT 15

US-09-252-991A-5260/c  
 ; Sequence 5260, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 5260  
 ; LENGTH: 2592  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-5260

Query Match 6.2%; Score 57.4; DB 4; Length 2592;  
 Best Local Similarity 53.3%; Pred. No. 6.6e-07;  
 Matches 121; Conservative 0; Mismatches 106; Indels 0; Gaps 0;  
 QY 35 TCGACGACCTAGCAGCTTTCATTTCAGTCGCTCAATCAGGCCACCTCACCGAAACTGCCG 94  
 Db 938 TCGACCTGCTCGGGTATTTCGAGTCGCGCGCGCCCACTCAGCTTCACCGCTGCCGCG 879  
 QY 95 AAAGATTAGGCATCCCGAGCCACACTTTCCAGACGAATCAGCCGAGTGGAAAAACAG 154  
 Db 878 AGGAAGTGGGGACCAACCCAGCTGCGGTTCAGCCAGCAGATCAACGCTGGAGAAAGAGC 819  
 QY 155 CAGGACCCCACTTTTCGACCGCGCGCGGCAAACTCGTCTCAACCAACGAGGCCAGC 214  
 Db 818 TGGCGACCCCTCTGTTCAGCGCGGTTTCATCGAGGCGATCGTCTCACCGACCGCGGCCAGC 759  
 QY 215 CCTTCTCAACACGCGCCATCGTCGAGAAATTCAACTCCGCC 261  
 Db 758 TTCTGCTCCGCGGCGGCTCGAGACGATCGATGCCGC 712

Search completed: April 13, 2005, 02:03:01  
 Job time : 209 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2005, 23:46:39 ; Search time 599 Seconds  
(without alignments)  
9408.686 Million cell updates/sec

Title: US-09-826-909-1\_COPY\_232\_1161

Perfect score: 930

Sequence: 1 gtgaccatgggcaacagcgg.....sgtatgcattagaagagggc 930

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5615251 seqs, 3030001701 residues

Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length  | ID | Description        |
|------------|-------|-------------|---------|----|--------------------|
| 1          | 930   | 100.0       | 1364    | 9  | US-09-826-909-1    |
| 2          | 930   | 100.0       | 3309400 | 9  | US-09-738-626-1    |
| 3          | 924   | 99.4        | 924     | 9  | US-09-738-626-3228 |
| 4          | 439   | 47.2        | 439     | 9  | US-09-826-909-3    |
| 5          | 129.6 | 13.9        | 77536   | 10 | US-09-940-316B-1   |
| 6          | 110.4 | 11.9        | 951     | 15 | US-10-156-761-3347 |
| 7          | 110.4 | 11.9        | 9025608 | 15 | US-10-156-761-1    |
| 8          | 99.8  | 10.7        | 9025608 | 15 | US-10-156-761-1    |
| 9          | 97.8  | 10.5        | 936     | 15 | US-10-156-761-1506 |
| 10         | 74.6  | 8.0         | 936     | 15 | US-10-156-761-4774 |
| 11         | 71.4  | 7.7         | 891     | 15 | US-10-156-761-7285 |

|    |      |     |         |    |                      |                   |
|----|------|-----|---------|----|----------------------|-------------------|
| 12 | 62.6 | 6.7 | 894     | 15 | US-10-156-761-4980   | Sequence 4980, Ap |
| 13 | 60.4 | 6.5 | 1008    | 15 | US-10-156-761-3218   | Sequence 3218, Ap |
| 14 | 60.2 | 6.5 | 876     | 15 | US-10-156-761-1152   | Sequence 1152, Ap |
| 15 | 60   | 6.5 | 891     | 17 | US-10-282-122A-31440 | Sequence 31440, A |
| 16 | 59.8 | 6.4 | 954     | 17 | US-10-282-122A-14715 | Sequence 14715, A |
| 17 | 59.6 | 6.4 | 987     | 15 | US-10-156-761-2141   | Sequence 2141, Ap |
| 18 | 57.8 | 6.2 | 882     | 9  | US-09-738-626-473    | Sequence 473, App |
| 19 | 57.6 | 6.2 | 1675    | 9  | US-09-938-641-1      | Sequence 1, Appli |
| 20 | 57.6 | 6.2 | 3309400 | 9  | US-09-738-626-1      | Sequence 1, Appli |
| 21 | 57.2 | 6.2 | 31248   | 19 | US-10-485-710-1      | Sequence 1, Appli |
| 22 | 57.2 | 6.2 | 35359   | 19 | US-10-485-710-2      | Sequence 2, Appli |
| 23 | 56.6 | 6.1 | 981     | 9  | US-09-738-626-2114   | Sequence 2114, Ap |
| 24 | 56   | 6.0 | 888     | 17 | US-10-398-221-2042   | Sequence 2042, Ap |
| 25 | 56   | 6.0 | 319630  | 17 | US-10-398-221-7      | Sequence 7, Appli |
| 26 | 56   | 6.0 | 3011208 | 17 | US-10-398-221-2058   | Sequence 2058, Ap |
| 27 | 54.4 | 5.8 | 912     | 15 | US-10-156-761-761    | Sequence 761, App |
| 28 | 54.4 | 5.8 | 1109    | 17 | US-10-398-221-1521   | Sequence 1521, Ap |
| 29 | 54.4 | 5.8 | 3259    | 17 | US-10-398-221-3552   | Sequence 3552, Ap |
| 30 | 53.4 | 5.7 | 1027    | 17 | US-10-282-122A-24068 | Sequence 24068, A |
| 31 | 52.4 | 5.6 | 986     | 17 | US-10-282-122A-13903 | Sequence 13903, A |
| 32 | 52   | 5.6 | 897     | 9  | US-09-815-242-4087   | Sequence 4087, Ap |
| 33 | 52   | 5.6 | 897     | 17 | US-10-282-122A-7352  | Sequence 7352, Ap |
| 34 | 51.8 | 5.6 | 3390    | 9  | US-09-767-878-1      | Sequence 1, Appli |
| 35 | 51.2 | 5.5 | 1830121 | 17 | US-10-329-670-1      | Sequence 1, Appli |
| 36 | 51.2 | 5.5 | 1830121 | 18 | US-10-158-865-1      | Sequence 1, Appli |
| 37 | 50.2 | 5.4 | 80161   | 17 | US-10-329-148A-1     | Sequence 1, Appli |
| 38 | 49.6 | 5.3 | 867     | 17 | US-10-282-122A-15044 | Sequence 15044, A |
| 39 | 49.4 | 5.3 | 762     | 9  | US-09-738-626-249    | Sequence 249, App |
| 40 | 49.4 | 5.3 | 892     | 19 | US-10-495-066-29     | Sequence 29, Appl |
| 41 | 49   | 5.3 | 2731748 | 18 | US-10-297-465A-1     | Sequence 1, Appli |
| 42 | 48.8 | 5.2 | 312     | 9  | US-09-974-300-1491   | Sequence 1491, Ap |
| 43 | 48.2 | 5.2 | 30000   | 11 | US-09-980-217-1      | Sequence 1, Appli |
| 44 | 46.8 | 5.0 | 909     | 9  | US-09-738-626-676    | Sequence 676, App |
| 45 | 46.8 | 5.0 | 1039    | 18 | US-10-494-672-53     | Sequence 53, Appl |

ALIGNMENTS

RESULT 1

US-09-826-909-1  
; Sequence 1, Application US/09826909  
; Patent No. US20020081674A1  
; GENERAL INFORMATION:  
; APPLICANT: MOECKEL, BETTINA  
; APPLICANT: FARWICK, MIKE  
; APPLICANT: HERMANN, THOMAS  
; APPLICANT: KREUTZER, CAROLINE  
; APPLICANT: PFEFFERLE, WALTER  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES FOR ENCODING OF THE LYSR2-GENE  
; FILE REFERENCE: 205551USOX  
; CURRENT APPLICATION NUMBER: US/09/826,909  
; CURRENT FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: DE 100 39 047.1  
; PRIOR FILING DATE: 2000-08-10  
; PRIOR APPLICATION NUMBER: DE 100 10 346.8  
; PRIOR FILING DATE: 2001-03-03  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (232)..(1161)  
; OTHER INFORMATION:  
US-09-826-909-1

Query Match 100.0%; Score 930; DB 9; Length 1364;  
Best Local Similarity 100.0%; Pred.No. 9.9e-293;  
Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GTGACCATGGGCAACGACGGCGGAGACTCGGAATCGACGACTACGAGCTTCAATTTCA 60
Db |||
QY 232 GTGACCATGGGCAACGACGGCGGAGACTCGGAATCGACGACTACGAGCTTCAATTTCA 291
Db |||
QY 61 GTGCTCAATCAGGCGACCTCAACGAACTGCGGAAAGATTAGGATCCCGAGGCCACA 120
Db |||
QY 292 GTGCTCAATCAGGCGACCTCAACGAACTGCGGAAAGATTAGGATCCCGAGGCCACA 351
Db |||
QY 121 CTTTCCAGAGATCAGCGAGTGGGAAACAGGAGGACCCGCTTTTCAGCGGCC 180
Db |||
QY 352 CTTTCCAGAGATCAGCGAGTGGGAAACAGGAGGACCCGCTTTTCAGCGGCC 411
Db |||
QY 181 GCGCGCAAACTCGTCTCAACCAACGAGGCCAGCTTCTCAACAGCGGAGCGCATC 240
Db |||
QY 412 GCGCGCAAACTCGTCTCAACCAACGAGGCCAGCTTCTCAACAGCGGAGCGCATC 471
Db |||
QY 241 GTGCGAGAATTCAACTCGCGCCCAACTGAAATCAAACTGCTCATGAGCCAGGAAAGGC 300
Db |||
QY 472 GTGCGAGAATTCAACTCGCGCCCAACTGAAATCAAACTGCTCATGAGCCAGGAAAGGC 531
Db |||
QY 301 ACAATCCGACTGACTTCAATGATCTTCTTGGGCACTTGGATGTCCTCGAATTCGGA 360
Db |||
QY 532 ACAATCCGACTGACTTCAATGATCTTCTTGGGCACTTGGATGTCCTCGAATTCGGA 591
Db |||
QY 361 ACAATCCGCGCGCAACACACCCCAAGTAGAATTCCTCACTCCACCAAGCGGCAATGCTC 420
Db |||
QY 592 ACAATCCGCGCGCAACACACCCCAAGTAGAATTCCTCACTCCACCAAGCGGCAATGCTC 651
Db |||
QY 421 CTGGTAGATCGTGTGTTTGGCTGATGAACTGCACTCGCAATGATTGTTGGCCCAAACTGCC 480
Db |||
QY 652 CTGGTAGATCGTGTGTTTGGCTGATGAACTGCACTCGCAATGATTGTTGGCCCAAACTGCC 711
Db |||
QY 481 GAGGTGTTGACTCTTTAGGTTGGGCGCACTGCTTCGTCAACGACTTGCCTAGCTGT 540
Db |||
QY 712 GAGGTGTTGACTCTTTAGGTTGGGCGCACTGCTTCGTCAACGACTTGCCTAGCTGT 771
Db |||
QY 541 CCGCGAGATCACCGGCTTGCCTCTCTTTCTGCGCAAGGAGAAATGCGCTTGATTACTGCG 600
Db |||
QY 772 CCGCGAGATCACCGGCTTGCCTCTCTTTCTGCGCAAGGAGAAATGCGCTTGATTACTGCG 831
Db |||
QY 601 GCGGAAGAACCTTTCTGTCGAGATGCGAGAGGTTTCGCGACCCGACTCTCTCATGATGCA 660
Db |||
QY 832 GCGGAAGAACCTTTCTGTCGAGATGCGAGAGGTTTCGCGACCCGACTCTCTCATGATGCA 891
Db |||
QY 661 TTAGCCGAGAGACCGGTTTCTGTCGAGATGCGAGAGGTTTCGAGTCCATGCACTCACACC 720
Db |||
QY 892 TTAGCCGAGAGACCGGTTTCTGTCGAGATGCGAGAGGTTTCGAGTCCATGCACTCACACC 951
Db |||
QY 721 GTGCGAGGCTTGTGAGCGAGGCTCTCGCGGTTGTTGTTGTTCCGATGGATGATCCGTAC 780
Db |||
QY 952 GTGCGAGGCTTGTGAGCGAGGCTCTCGCGGTTGTTGTTGTTCCGATGGATGATCCGTAC 1011
Db |||
QY 781 CTTCCACAGTGGGAATCGTCAACGCGCCACTTAGTCCACCGCTTTATAGGGAATAGGT 840
Db |||
QY 1012 CTTCCACAGTGGGAATCGTCAACGCGCCACTTAGTCCACCGCTTTATAGGGAATAGGT 1071
Db |||
QY 841 TTGGTGTGCGCACTCAACGCGGGCGCGCACTGCGGTGGATTAACCTTCGGAAGTTCGTG 900
Db |||
QY 1072 TTGGTGTGCGCACTCAACGCGGGCGCGCACTGCGGTGGATTAACCTTCGGAAGTTCGTG 1131
Db |||
QY 901 GCGGGATCGAGGTATGCAATTAGAGAGGCG 930
Db |||
QY 1132 GCGGGATCGAGGTATGCAATTAGAGAGGCG 1161
Db |||
```

## RESULT 2

US-09-738-626-1/c

; Sequence 1, Application US/09738626

; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

```
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1
```

Query Match 100.0%; Score 930; DB 9; Length 3309400;  
Best Local Similarity 100.0%; Pred. No. 1.7e-291;

Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GTGACCATGGGCAACGACGGCGGAGACTCGGAATCGACGACTACGAGCTTCAATTTCA 60
Db 3128423 GTGACCATGGGCAACGACGGCGGAGACTCGGAATCGACGACTACGAGCTTCAATTTCA 3128364
QY 61 GTGCTCAATCAGGCGACCTCAACGAACTGCGGAAAGATTAGGATCCCGAGGCCACA 120
Db 3128363 GTGCTCAATCAGGCGACCTCAACGAACTGCGGAAAGATTAGGATCCCGAGGCCACA 3128304
QY 121 CTTTCCAGAGATCAGCGAGTGGGAAACAGGAGGACCCGCTTTTCAGCGGCC 180
Db 3128303 CTTTCCAGAGATCAGCGAGTGGGAAACAGGAGGACCCGCTTTTCAGCGGCC 3128244
QY 181 GCGCGCAAACTCGTCTCAACCAAGAGGCGCACCGCTTCTCAACAGCGCGGCCATC 240
Db 3128243 GCGCGCAAACTCGTCTCAACCAAGAGGCGCACCGCTTCTCAACAGCGCGGCCATC 3128184
QY 241 GTGCGAGAATTCAACTCGCGGCAACTGAAATCAAACTGCTCATGAGCCAGGAAAGGC 300
Db 3128183 GTGCGAGAATTCAACTCGCGGCAACTGAAATCAAACTGCTCATGAGCCAGGAAAGGC 3128124
QY 301 ACAATCCGACTGAGCTTCAATGCTTCTTGGGCACTTGGATGTCCTCGGACTTATCCGA 360
Db 3128123 ACAATCCGACTGAGCTTCAATGCTTCTTGGGCACTTGGATGTCCTCGGAACTTATCCGA 3128064
QY 361 ACATTCGCGCGCGCAACACCCCAAGTAGAATTCCTCAACAGCGCGGAGCAATGCTC 420
Db 3128063 ACATTCGCGCGCGCAACACCCCAAGTAGAATTCCTCAACAGCGCGGAGCAATGCTC 3128004
QY 421 CTGTTAGATCGTGTGTTTGGCTGATGAACTGCACTCGCATTTAGTTGGCCCAACCTGTC 480
Db 3128003 CTGTTAGATCGTGTGTTTGGCTGATGAACTGCACTCGCATTTAGTTGGCCCAACCTGTC 3127944
QY 481 GAGGTGTTGACTCTTTAGGTTGGGCGCACTGCTTCGTCAACGACTTGCCTAGCTGT 540
Db 3127943 GAGGTGTTGACTCTTTAGGTTGGGCGCACTGCTTCGTCAACGACTTGCCTAGCTGT 3127884
QY 541 CCGCGAGATCACCGGCTTGCCTCTCTTTCTGGCAAGGAGAAATGCGCTTGATTACTGCG 600
Db 3127883 CCGCGAGATCACCGGCTTGCCTCTCTTTCTGGCAAGGAGAAATGCGCTTGATTACTGCG 3127824
QY 601 GCGGAAGAACCTTTCTGTCGAGATGCGAGAGGTTTCGCGACCCGACTCTCTCATGATGCA 660
Db 3127823 GCGGAAGAACCTTTCTGTCGAGATGCGAGAGGTTTCGCGACCCGACTCTCTCATGATGCA 3127764
```



QY 661 TTAGCGAAGAGCCGGTTTTGTTCCAAATGTGGTTTTTCGAATCCATGGAATCACCACC 720  
DB 3127763 TTAGCGAAGAGCCGGTTTTGTTCCAAATGTGGTTTTTCGAATCCATGGAATCACCACC 3127704  
QY 721 GTCCGAGGCTTGTGACGCGCAGTCTCGGCGTTGGTGTGGTCCGATGGATGATCCGTAC 780  
DB 3127703 GTCCGAGGCTTGTGACGCGCAGTCTCGGCGTTGGTGTGGTCCGATGGATGATCCGTAC 3127644  
QY 781 CTTCCACAGTGGGAATCGTCAAGCCCACTTAGTCCACCCGCTTTATAGGGAACATAGGT 840  
DB 3127643 CTTCCACAGTGGGAATCGTCAAGCCCACTTAGTCCACCCGCTTTATAGGGAACATAGGT 3127584  
QY 841 TTGCTGTGCGCACTCAACCGCGGCGCGCACCTCGCGTGGATTAACCTTCCGGAAGTTGCTG 900  
DB 3127583 TTGCTGTGCGCACTCAACCGCGGCGCGCACCTCGCGTGGATTAACCTTCCGGAAGTTGCTG 3127524  
QY 901 GCGGATCGAGGTATGCAATAGAGAGGCG 930  
DB 3127523 GCGGATCGAGGTATGCAATAGAGAGGCG 3127494

RESULT 3

US-09-738-626-3228  
; Sequence 3228, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOHO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1998-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 3228  
; LENGTH: 924  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3228

Query Match 99.4%; Score 924; DB 9; Length 924;  
Best Local Similarity 100.0%; Pred. No. 7.8e-291;  
Matches 924; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 ATGGCGAAGAGCGGGAGACCTGCGAATCGAGCCTACGAGCTTCATTTCACTCGCT 66  
DB 1 ATGGCGAAGAGCGGGAGACCTGCGAATCGAGCCTACGAGCTTCATTTCACTCGCT 60  
QY 67 CAATCAGGCGCCTCACAGAACTGCGAAAGATTAGGCATCCCGCAGCCACACTTTCC 126  
DB 61 CAATCAGGCGCCTCACAGAACTGCGAAGATTAGGCATCCCGCAGCCACACTTTCC 120  
QY 127 AGACGAATCAGCGAGTGGAAAAACACGAGGACCCCACTTTTCGACCGCGCGCGCGC 186  
DB 121 AGACGAATCAGCGAGTGGAAAAACACGAGGACCCCACTTTTCGACCGCGCGCGCGC 180  
QY 187 AAATCGTCTCAACCAAGGAGCCACGCTTCCTCAACGAGCGGCGCATCTGCGCA 246  
DB 181 AAATCGTCTCAACCAAGGAGCCACGCTTCCTCAACGAGCGGCGCATCTGCGCA 240

RESULT 4

US-09-826-909-3  
; Sequence 3, Application US/09826909  
; Patent No. US20020081674A1  
; GENERAL INFORMATION:  
; APPLICANT: MOECKEL, BETTINA  
; APPLICANT: FARWICK, MIKE  
; APPLICANT: HERMANN, THOMAS  
; APPLICANT: KREUTZER, CAROLINE  
; APPLICANT: PFEFFERLE, WALTER  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES FOR ENCODING OF THE LYSR2-GENE  
; FILE REFERENCE: 205551US0X  
; CURRENT APPLICATION NUMBER: US/09/826,909  
; CURRENT FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: DE 100 39 047.1  
; PRIOR FILING DATE: 2000-08-10  
; PRIOR APPLICATION NUMBER: DE 100 10 346.8  
; PRIOR FILING DATE: 2001-03-03  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 439  
; TYPE: DNA

QY 247 GAAATTCATCCGCGCAACTGAAATCAAAACGCTCATGGACCCAGAAAAAGGCAATC 306  
DB 241 GAAATTCATCCGCGCGCAACTGAAATCAAAACGCTCATGGACCCAGAAAAAGGCAATC 300  
QY 307 CGATCGGACTTCATGCAATCTTGGGCACTTGGATGGTCCCGAACTTATCCGAACATTC 366  
DB 301 CGATCGGACTTCATGCAATCTTGGGCACTTGGATGGTCCCGAACTTATCCGAACATTC 360  
QY 367 CGCGCGCAACACCCCAAGTGAATTCACCTCCACCAAGCGGCAAGTGTCTCTGGTA 426  
DB 361 CGCGCGCAACACCCCAAGTGAATTCACCTCCACCAAGCGGCAAGTGTCTCTGGTA 420  
QY 427 GATCGTGTGTTGGCTGATGAACCTGACCTCGCATTTAGTTGGCCCCCAACCTCCCGAGTT 486  
DB 421 GATCGTGTGTTGGCTGATGAACCTGACCTCGCATTTAGTTGGCCCCCAACCTCCCGAGTT 480  
QY 487 GGTACTCTTTAGGTTGGCGCACTGCTTGTGTAACGACTTGCCTAGCTGTTCCTCGCA 546  
DB 481 GGTACTCTTTAGGTTGGCGCACTGCTTGTGTAACGACTTGCCTAGCTGTTCCTCGCA 540  
QY 547 GATCACCGGCTTGCCCTCTTTCTGGCCAAAGGAGAAATTCGCTTCACTTCGCGCGAA 606  
DB 541 GATCACCGGCTTGCCCTCTTTCTGGCCAAAGGAGAAATTCGCTTCACTTCGCGCGAA 600  
QY 607 GAACTTTCTGTCGGGATCGAGCAGTTTCGGCACCCGACTCCTCATGATGATTCATTC 666  
DB 601 GAACTTTCTGTCGGGATCGAGCAGTTTCGGCACCCGACTCCTCATGATGATTCATTC 660  
QY 667 GAAGAAGCGGTTTGTCCCAATGTGTTTTCGAATCGATGGAACCTCACACCGTCGCA 726  
DB 661 GAAGAAGCGGTTTGTCCCAATGTGTTTTCGAATCGATGGAACCTCACACCGTCGCA 720  
QY 727 GGGCTTGTGACGCGAGGTCGCGGCTTGGTTCGATGGATGATTCGCTTCCCTCC 786  
DB 721 GGGCTTGTGACGCGAGGTCGCGGCTTGGTTCGATGGATGATTCGCTTCCCTCC 780  
QY 787 ACAGTGGGAATCGTGCAACGCGCCACTTAGTCCACCCGCTTATAGGGAACCTAGGTTTGGTG 846  
DB 781 ACAGTGGGAATCGTGCAACGCGCCACTTAGTCCACCCGCTTATAGGGAACCTAGGTTTGGTG 840  
QY 847 TGGCGACTCAACGCGGCGCGGCACTCGCGTGGATTAACCTTCGGAAGTTCTGCGCGGA 906  
DB 841 TGGCGACTCAACGCGGCGCGGCACTCGCGTGGATTAACCTTCGGAAGTTCTGCGCGGA 900  
QY 907 TCGAGGTATGCAATAGAGAGGCG 930  
DB 901 TCGAGGTATGCAATAGAGAGGCG 924

```
; ORGANISM: Corynebacterium glutamicum
US-09-826-909-3

Query Match      47.2%; Score 439; DB 9; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.5e-132;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 CCATCGTCGACGAATTCAACTCCGCGCAACTGAATCAAAACGCTCATGGACCCAGAAA 295
Db 1 CCATCGTCGACGAATTCAACTCCGCGCAACTGAATCAAAACGCTCATGGACCCAGAAA 60

QY 296 AAGGCACATCCGACCTGGACTTCATGCAATTCCTTGGGCACTTGGATGTTCCCGGAACCTTA 355
Db 61 AAGGCACATCCGACCTGGACTTCATGCAATTCCTTGGGCACTTGGATGTTCCCGGAACCTTA 120

QY 356 TCCGAACATCCGCGCGGACACCCCAAGCTAGAAATTCCACTCACAAGCGGAGCAA 415
Db 121 TCCGAACATCCGCGCGGACACCCCAAGCTAGAAATTCCACTCACAAGCGGAGCAA 180

QY 416 TGCTCTCTGTAGATCGTGTGTTGGCTGATGAACCTGACTGCTAGTTGGSCCCCAAC 475
Db 181 TGCTCTCTGTAGATCGTGTGTTGGCTGATGAACCTGACTGCTAGTTGGSCCCCAAC 240

QY 476 CTGCCGAGTGGTACCTCTTTAGGTTGGGCGCCACTGCTTGGTCAACGACTTGCCTTAG 535
Db 241 CTGCCGAGTGGTACCTCTTTAGGTTGGGCGCCACTGCTTGGTCAACGACTTGCCTTAG 300

QY 536 CTGTTCCCGCAGATCACCGGCTTGCTCTCTTTCTGGCGAAGGAAATTCGGTTGATTA 595
Db 301 CTGTTCCCGCAGATCACCGGCTTGCTCTCTTTCTGGCGAAGGAAATTCGGTTGATTA 360

QY 596 CTGCGCGGGAAGAACTTTCGTGGCGATGCGAGCAGGTTTGGCACCCGACTCCTCATGG 655
Db 361 CTGCGCGGGAAGAACTTTCGTGGCGATGCGAGCAGGTTTGGCACCCGACTCCTCATGG 420

QY 656 ATGCATTAGCCGAAGAAGC 674
Db 421 ATGCATTAGCCGAAGAAGC 439

RESULT 5
US-09-940-316B-1/c
; Sequence 1, Application US/09940316B
; Publication No. US20030175901A1
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
; FILE REFERENCE: 30062-20026.11
; CURRENT APPLICATION NUMBER: US/09/940,316B
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/410,551
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 77536
; TYPE: DNA
; ORGANISM: Streptomyces hygroscopicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52275) ... (71465)
```

```
US-09-940-316B-1

Query Match      13.9%; Score 129.6; DB 10; Length 77536;
Best Local Similarity 50.7%; Pred. No. 1.1e-30;
Matches 374; Conservative 0; Mismatches 349; Indels 15; Gaps 2;

QY 27 CCTCGAATCGACGACCTTACGACGCTTCATTTAGTCGCTCAATTCAGCGCCACCTCACCGA 86
Db 5489 CATGAGATACGGAGATTGCAAGTGTTCACGGCTTGGCCGAGCAGCAGACGATGACCAT 5430

QY 87 AACTGCCGAAAGATTAGGCATCCGCGACCCACACTTTTCCAGAGCAATTCAGCCGAGTGA 146
Db 5429 CACCGGGAACGGCTCAACATCTCCAGCGACACTGTCCAGGGCCATCCGGGCGCTGA 5370

QY 147 AAAACACGCGACGCCCACTTTTTCGACCGCGCGCGCGCAAACTGCTCTCAACCAACG 206
Db 5369 CGCAGATGGGCGGTGCGGCTGTTTCGACCGGCAACCGGCTCCGGCTCAACAGTA 5310

QY 207 AGGCCACGCTTCTCAACACGCGCGCATCGTGCAGAAATTCAACTCCGCGCAAC 266
Db 5309 CGCGAGTCTTTCGCGCCCAAGCCCTCGCGCGCATAGTAGGTTCTCCGGGCGCAGCA 5250

QY 267 TGAATCAAAACGCTCATGACCCAGAAAAAGGCACAATCCGACTTGTGACTTTCATGCTTC 326
Db 5249 GCGCATCACGCCCTGATCGACCCCGACACCGGCAACCGTGGCACTGGGTTTCTGCACTC 5190

QY 327 CTTGGGCACTTGGATGTTCCCGGAACCTTATCCGAACATTCGCGCGGCAACACCCCAAGT 386
Db 5189 CTACGGCACCTGGTGGTGGTCCCGGACCTCTCTCGCGGGTACCGCGGCTGGCTCCGGCAC 5130

QY 387 AGAATTCCAACTCCACCAAGCGGAGCAATTCCTCTCTGTAGATCGTGTGTTGGCTGATGA 446
Db 5129 CACCTTCGAATCGCGCGGAGCGCGCGCGCGCTGGTGCAGCAGCTCCGCAACAGCAG 5070

QY 447 AACTGACCTCGCATTAGTTGGCCCCCAAACTTCGCGAGGTTGGTACTCTTTAGGGTGGGC 506
Db 5069 GCTGGACCTGGGGCTGACCGACGACGACCGCGCGGG-----ACGATGTGGAGTGGAC 5016

QY 507 GCGACTGCTTCGTCACGACTTGGCTAGCTGTTCGCGCAGATCACCGGCTTGGCTCCTT 566
Db 5015 ACCACTCGGGAAGAGGACCTGTGCTCTGCTCCACCGCGGACACCGGCTGGGC----- 4961

QY 567 TTCTGGCCAAAGGAGAAATTCGCGTTGATTAATTCGCGCGGAGAAACCTTTTCGTGGGATCGC 626
Db 4960 ----AGACGCGCGCGGGTGGGACCGCGGAGCTGGCCGATGAGGGTTCGTGGCACTGA 4905

QY 627 AGCAGGTTTGGCACCCGACTCTCTCATGATGATAGCCGAAAGACCGGTTTGTGTTCC 686
Db 4904 GCCGCTCTTCGGAATGGCGCCAGATCACGACCGGCTGTGCGCGCGCGGCGGATTCGTGCC 4845

QY 687 CAATGTGTTTTCGAATCCATGGAACCTCACCGCTCGCAGGGCTTGTACGCGGAGGTCT 746
Db 4844 GCGGTCGCTCTCGAGAGCACCGAGCTGAGCACGCTGGGGGCCCTTGGTCGCGCGCGGCT 4785

QY 747 CGGCTTGGTGTGTTTC 764
Db 4784 CGGTGTGGCTCGCCCC 4767

RESULT 6
US-10-156-761-3347
; Sequence 3347, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
```

```
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3347
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(951)
US-10-156-761-3347

Query Match      11.9%; Score 110.4; DB 15; Length 951;
Best Local Similarity 48.4%; Pred. No. 4.2e-25;
Matches 411; Conservative 0; Mismatches 421; Indels 18; Gaps 3;

QY 51 CTTCAATTTTCAGTCGCTCAATCAGGCCACTCACCAGAACTGCGGAAAGATTAGGATCC 110
DB 90 CTTCCCGGGGTGGCCGACGAGCAGTCAACCGGGCGCGGAGGAGATGAGTCC 149
QY 111 CGAGCCCACTTTCCAGACGAAATCAGCGGAGTGGAAGAAACAGCGAGGACCCCACTTTT 170
DB 150 CGAGTCCACCTCTCCCGCTCCCTCGCTCGGCTCGAGCAGGACCTGGGCGTGCACCTTT 209
QY 171 CGACCGCGCGCGCGCAACTCGTCTCAACCAACAGGCGCAGCTTCTTCAACACGCG 230
DB 210 CGCCCGCGCGCGCGCAGCTCGCTCACCCCGCGCGCGCTACGTTCTTCCCTCCGT 269
QY 231 CAGCGCCCATCTGTCGAGAAATTCAACTCCGCGCAACTGAAATCAAAACGCTCATGGAACC 290
DB 270 CGAGCGCGCTCGCGAGATCGAGGGCGCGCGAGGAGTGGCGCGGACCGACCC 329
QY 291 AGAAAAAGCACAATTCGAGCTGGAATTCATGATTTCTTTGGGCACTTGGATGGTCCCGA 350
DB 330 GGCACCGCGAAGTTCGCTTCGGCTTTCTGCACACCACTGGGCTCGAGACCGTGC 389
QY 351 ACTTATCCGAATTCGCGCGCGACACCCCAACCTAGAAATTCACCACTCCACCAAGCGCG 410
DB 390 GCTGATCCGAGCTTCGCGCGCGATCATCGCGCGTTCGCTTCAGCCTGCTCGAAGCTA 449
QY 411 AGCAATGCTCTGCTAGATGCTGTTTGGCTGATGAACTGACCTCGCATTTAGTTGGGCC 470
DB 450 CGCGGAGCGATGATCGAGCGCTCGGGCGGGGAGTTGGACCTCTGTCTGACCTCGC 509
QY 471 CAAACCTGCGAGTTGTGTAACCTTTTAGGGTGGGCGCACCTGCTTCTGTAACGACTTGC 530
DB 510 CGTGCAGGACGCGC-----CGACCTGGTGGCGCGCGCTCGACGAGCAGAACTGCG 563
QY 531 CTAAGTGTTCGCGAGATCAGCGGCTTGCCTCTTTCTGCGCAAGGAGAAATTCGCTT 590
DB 564 GCTGCTGCTCGCGCGCGACACCGCTCGCGCAACGCAAGCGC-----GTGCGCT 614
QY 591 GATTACTGCGGGAAGAACCTTTCTGTGCGATGCGAGCAGTTTCGGCACCGCACTCT 650
DB 615 CGCGGAGCGCGGAGACCTTCTGTGACCTCGAAACCGGTTAGGGCTGCGCGGAT 674
QY 651 CATGATGATTTAGCGGAGAGCGGTTTGTTCCTCAATGTTGTTTTCGAATTCATGGA 710
DB 675 CACCGACGAGCTGCGCAGGAGCGGTTTTCAGACGCGGCATCGCTTCGAGGCGGAG 734
QY 711 ACTCACCACCTGCGAGGCTTGTACGCGCAGGTCTCGCGCTTGTGTTGTTCCGATGA 770
DB 735 GCGGAGAGCTACGCGGCTTGTGCGCGCGGCTCGGGGTTCGCGCTCTCTCGCGCAC 794
QY 771 TCACTCGTACCTTCCACAGTGGAACTCGTGAACGCCCACTTAGTCCACCGCTTATAG 830
DB 795 GCGCGTCCCGCGGAGTTG---TCGAGCTGACGCTACGCGCGCGCGCGGTCG 851
QY 831 GGAACTAGTTTGGTGGCAGCTCAACCGCGGGCGCGACCTGCGGTGGATAACTTCG 890
```

```
DB 852 CGAGATCGGCTCGCTCGCTGAGACACCGGACACACCGCGGTGGCGCTTCAA 911
QY 891 GAAGTTCTGTG 900
DB 912 GAAGTTCTGTG 921

RESULT 7
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PENDING FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match      11.9%; Score 110.4; DB 15; Length 9025608;
Best Local Similarity 48.4%; Pred. No. 1.2e-23;
Matches 411; Conservative 0; Mismatches 421; Indels 18; Gaps 3;

QY 51 CTTCAATTTTCAGTCGCTCAATCAGGCCACTCACCAGAACTGCGGAAAGATTAGGATCC 110
DB 4176160 CTTCCCGGGGTGGCCGACGAGCAGTCAACCGGGCGCGGAGGAGATGAGTCC 4176101
QY 111 CGAGCCCACTTTCCAGACGAAATCAGCGGAGTGAAGAAACAGCGAGCACCCTTTT 170
DB 4176100 CGAGTCCACCTCTCTCCGCTCCCTCGCTCGAGGAGACCTGGGCGTTCGACCTCT 4176041
QY 171 CGACCGCGCGCGCAAACTCGTCTCAACCAACGAGGCCACGCTTCTTCAACACCGC 230
DB 4176040 CGCCCGCGCGCGCGCAGCTCTCGCTCACCCCGCGCGCGTCTGCTTCTCGCTCCGT 4175981
QY 231 CAGCGCCCATCTGTCGAGAAATTCAACTCCGCGCAACTGAAATCAAAACGCTCATGACCC 290
DB 4175980 CGAGCGCGCTCGCGGAGATCGAGCGGCGCGCGAGGAGTGGCGCGGACCGC 4175921
QY 291 AGAAAAAGGCAATTCGACTGGACTTCATGATTTCTTGGGCACTTGGATGGTCCCGA 350
DB 4175920 GGCACCGCGCAAGTTCGCTTTCGCTTTCGACACCACTGGGCTCGAGACCGTGC 4175861
QY 351 ACTTATCCGAACATTCGCGCGCAACCCCAACCTAGAAATTCCAACTCCACCAAGCGCG 410
DB 4175860 GCTGATCCGAGCTTCCGCGCGCATCTCCGCGCGCTCGCTTCAGCTCGTCAGAACTA 4175801
QY 411 AGCAATGCTCTGCTGATGCTGTTTGGCTGATGAACTGACCTCGCATTTAGTTGGGCC 470
DB 4175800 CGCGGAGGAGTATCGAGCGGCTCGGCGCGGAGTTGGACCTCTGTCTGACCTCGC 4175741
QY 471 CAAACCTGCGAGGTTGTTGTAACCTTTTAGGGTGGGCGCACCTGCTTCGTCAGGACTTGC 530
```

Db 4175740 CGTGGCGGACGGC-----CCGACTGTGTGGCCCGCGCTCGACGAGCAGAGCTGCG 4175687  
QY 531 CTTAGCTGTTCCCGCAGATCACGGGCTTGCCTCTCTTTCTGGCCAAAGAGAAATTCGCCGTT 590  
Db 4175686 GCTGTGTGTCGGGCGGACCGGCTCGCGCAGCAAGCGC-----GTGCGCT 4175636  
QY 591 GATTACTGCGGCGGAAGAACCTTTCTGTGGCGATGCGAGCAGGTTTTCGGCACCAGGACTCCT 650  
Db 4175635 CGCCGAGCGCGCAGAGACTTTCGTGACCTCGAAACCGGTTACGGGCTGCGCGCTAT 4175576  
QY 651 CATGATGATTAAGCCGGAAGACCGGTTTTCCTCAATGTTGTTTCGAATTCATGGA 710  
Db 4175575 CACCAGCAGCTCTCCAGGAGCGCGTTTCAGACCGGCACTCGCTTCGAGGCGGAGGA 4175516  
QY 711 ACTCACCACCTGCGAGGGCTGTGAGCGAGGCTCGGCGTTGTGTGTTTCGATGGA 770  
Db 4175515 GCGGAGACGCTACCGGGCGCTGTGCGCGCGGGCTCGGGGTGCGCTCTTCGCGCCACC 4175456  
QY 771 TGATCCGTACTTTCACAGTGGGAATCGTCAACGCCCACTTAGTCCACCGCTTATAG 830  
Db 4175455 GCGCGTGGCCCGCGGAGTTG---TCGAGCTGACGGTCACGGCCCGCGGCGGTCCG 4175399  
QY 831 GAACTAGTTGTTGGTGGCACTCAACCGCGGGCGGCACTGCGGTGGATTAATTCG 890  
Db 4175398 CGAGATCGGCTGCGCTGGCTGGACGGCCACCGGACACACCGCGGTGGCGGCTTCAA 4175339  
QY 891 GAACTGCTG 900  
Db 4175338 GAACTGCTG 4175329

## RESULT 8

US-10-156-761-1

; Sequence 1, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 1  
; LENGTH: 9025608  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (418715)  
; OTHER INFORMATION: a, t, c, g, other or unknown

US-10-156-761-1

Query Match 10.7%; Score 99.8; DB 15; Length 9025608;  
Best Local Similarity 47.8%; Pred. No. 3.5e-20;  
Matches 364; Conservative 0; Mismatches 382; Indels 15; Gaps 2;  
QY 5 CCATGGGCAACGACGGCGGAGACCTTGGGAATCGACGACTACGAGCTTTCATTTTCAGTCG 64  
Db 1859954 CCATGAGGAGGAAGAAATCACGCTGACCTCCAGGTACTCCGCTCTTCAGCGCGGTG 1860013  
QY 65 CTCATCAGGCCACCTTACCGAACTCGCGAAGATTAGGATCCGCGAGCCACACTTT 124  
Db 1860014 CAGAGGGAACGACGCTGTGACCGCGCGGAGCGCCCGTATCACCCAGCGCCGCTGT 1860073

QY 125 CCAGACGAAATCAGCGGAGTGAATAAACACGACGAGCAACCCCACTTTTCAGACGCGCCGCC 184  
Db 1860074 CGCGGGCCCTGAACCGGCTCGAGACGAGGTCGCGCGGGAATCTTTCAGCGCGTCGGAC 1860133  
QY 185 GCAAACTCGTCTCTCAACACGAGCGCACGCTTTCCTCAACACGACGCGCACTCGTCG 244  
Db 1860134 CGGTACTGCGCTCTACCCCGCGGACGGGTGTTCAAGGAGTACGTGACCGGGTGTCTG 1860193  
QY 245 CAGAATTAACCTCGCGCGCAACTGAAATCAAAACGCTCATGACCCGAAAAAGGACAA 304  
Db 1860194 ACAGCTACGACCGGGGCGGAGCGCTGCGGAGATCGTCAACCGGACGCGGTGCGG 1860253  
QY 305 TCCGACTGGAATTCATGATCTCTTGGGCACTTCGATGCTCCCGAACTTATCCGAAT 364  
Db 1860254 TGTGCTGGCTTCTCTGACACCTCGGCACTTCGCTGGTGGTCCCGAGGCTTGTACGAG 1860313  
QY 365 TCCGCGCGCAACACCCCAACGTAAGATTTCCAACTCCCAAGCGCGCAGCAATGCTCTCG 424  
Db 1860314 TCCGCGAGGCGTTCCCGCAGGCGCGGTTTCGAGCTGCAACGAGGACGCGGAGATCG 1860373  
QY 425 TAGATCGTGTGTTTGGCTGATGAATACTGACCTCGCATTAGTTGGCCCCAAACCTGCG 484  
Db 1860374 CCGGCACTCTCTCGACGCGCACCGCGGACCTGATCATCACGACGAGCGCGCCCA 1860431  
QY 485 TTGCTACTCTTTAGGGTGGGCGCACTGCTTCGTCAACGACTTGCCTAGCTGTTCCG 544  
Db 1860432 ----CCCCCTGATCACCTGCGCGGCTCTCTCTGTGGAGCCACTACGCGTGGCCG 1860487  
QY 545 CAGATCACCGGCTTGCCTCTCTTTCTGCGCAAGAGAAATTCGCCGTTGATTACTGCG 604  
Db 1860488 CCGGCAACCGGCTGCGC-----AGGCGGCGCGCGGCGCGCTGCGCGAGGTGCGCG 1860538  
QY 605 AAGAACCTTTCGTGGCGAGTGGACGAGGTTTCGGCACCGGACTCCTCATGATGATGAT 664  
Db 1860539 AGGAGCGGTTTCATCTCTCTCAAGCCGCGGTACGCGGCTGCGCGGATCAGCGAG 1860598  
QY 665 CCGAAGAAGCGGTTTGTTCCTCAATGTTGTTTTCGAATCCATGGAATCACCACCGTCG 724  
Db 1860599 GCCTCGAGGCGGCTTTCACCCGCTCTGCGCTTCGAGGGGAGAGGTTGAGACCTTGC 1860658  
QY 725 CAGGCTTGTGACGCGCAGGTCCTCGCGCTGTTGTTGTTCCG 765  
Db 1860659 GGGCGCTGTGTCGCGCGGCTCTGCGTGTGTTCCCTGATCCCG 1860699

## RESULT 9

US-10-156-761-1506  
; Sequence 1506, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 1506  
; LENGTH: 936  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: CDS

; LOCATION: (1)..(936)  
US-10-156-761-1506

Query Match 10.5%; Score 97.8; DB 15; Length 936;  
Best Local Similarity 47.7%; Pred. No. 5.5e-21;  
Matches 362; Conservative 0; Mismatches 382; Indels 15; Gaps 2;

QY 7 ATGGCAACGACGGCGGAGACTCGGAATCGACGACCTACGAGCTTCATTTAGTCGCT 66  
Db 1 ATGAGGAGAGAGATGCAAGCATGACCTCCAGGTACTCCGCTCTTCCAGGCGGTGCA 60

QY 67 CAATCAGGACCTCAGCGAACTCGCGAAGATTAGGATCCCGACGCCACATTTCC 126  
Db 61 GAGGAAACGACGGTACCGACCGCCGCGGAGCGCCGTATCACCGACCGCCCTGTG 120

QY 127 AGACGAATCAGCGAGTGGMAAAACACGAGGACCCCACTTTTCGACCGGCGCGCG 186  
Db 121 CGGCGCTTGAACCGCTCGAGACAGGTGCGGCGGGAATCTTCAGGGGTGCGAGC 180

QY 187 AAATCTGCTCTCAACCAAGAGGCGACCGCTTCTCAACCGCGAGCGCCATCGTCGA 246  
Db 181 GTACTGCGCTCACCCGCGGAGCGGTGTTCAAGGAGTACGTGACGCGGTGCTGAC 240

QY 247 GAATTAATCCGCGCACTGAAATCAAAAGCTATGAGACCCAGAAAAGGCAATC 306  
Db 241 AGCTACGACCGGCGCGGCGAGCCGTGCGGAGATCGTCGACCGCGAGCGCGGTG 300

QY 307 CGACTGGACTTCATGANTCTTGGGCACTTGGATGTTCCCGAACTTATCGAATTC 366  
Db 301 TCGTGGCTTCTGACACCTTGGGCACTTGGTGGTGGCGGAGTGTGACGAGTTTC 360

QY 367 CGCGCGCAACACCCAAAGTAGAATTCACACTCCAGCGGAGCAATGCTCTGTTA 426  
Db 361 CGGAGGCGTTCGCGAGCGCGGTTCGAGCTGCACGAGGCGGAGATCGACTGACC 420

QY 427 GATCGTGTGCTGATGAATCACTGACCTCGCATTTAGTTGGCCCAAGCTTCCGAGTT 486  
Db 421 CGGCACTCTCTGAGCGCACCGGACCTGATCATCACGAGCGAGCGCGCCCA--- 476

QY 487 GGTACTCTTTAGGTGGGCGCACTGCTTGTGTAACGACTTGCCTAGCTGTTCCGCA 546  
Db 477 ---CCCTGTATCACCTGGGCGGTCTCTCTGGTGGAGCACTACGGGTGGCGCTCCCGCC 534

QY 547 GATCACCGGCTGCTCTCTTTCTGGCCAAAGAGAAATTGCCGTGTGATTACTCGCGCGAA 606  
Db 535 CGGCAACCGCTGCGC-----AGGCGCGCGGCGCGTCTGGCGAGTGGCGGAG 585

QY 607 GAACCTTCTGGGAGTGGAGAGGTTTCGGACCCGACTCCTCATGATGATCATAGCC 666  
Db 586 GAGCGCTTCATCTGCTCAAGCCCGGTACGGGCTGCGGCGCATCAGCGAGACGTTGTGC 645

QY 667 GAAGAAGCGGTTTGTTCCTCAATGTTGTTTGAATCCATGAACTCACACCGTCGA 726  
Db 646 CTCGAGGCGGCTTACCCCGCTGCTGCGCTTCGAGGCGGAGGAGGAGGAGACCTTGGG 705

QY 727 GGGCTGTGTCAGCGAGGCTCTCGCGGTGTTGTTGGTTCCG 765  
Db 706 GGCCTGTGTCGCGCGGTCTGGTGTGTTCCCTGATCCCG 744

## RESULT 10

US-10-156-761-4774  
; Sequence 4774, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 4774  
; LENGTH: 936  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(936)  
US-10-156-761-4774

Query Match 8.0%; Score 74.6; DB 15; Length 936;  
Best Local Similarity 44.7%; Pred. No. 2.1e-13;  
Matches 342; Conservative 0; Mismatches 414; Indels 9; Gaps 1;

QY 5 CCATGGGCAACGACGGCGGAGACCTGCGAAATCGACGACTACGCAAGTTCATTTGAGTCG 64  
Db 17 CCATGAATCAGAACGATGGCGCGTGGATCTCCAGCAGATGGATACGTGTCGCGCTCG 76

QY 65 CTCATATCAGGCCACCTCACCGAAACTGCCGAAAGATTAGGCATCCGCGAGCCACACTTT 124  
Db 77 CGGAAACCGCAACTTACCCGCGCGGAGCGTGTCTCGTGGTGCAGTGTGCTCACTCA 136

QY 125 CAGACGAAATCAGCGGAGTGGAAAAACAGCAGGACCCCACTTTTCAGCCGCGGCGCC 184  
Db 137 GTCAACCGATCGCGGCTTGGAAACGAGGAGTCCGCGGTCAAGCTGTTCGCCCGATCCAG 196

QY 185 GCAAACTCGTCTCAACCAACGAGGCGCACGCTTCTCAACACGAGCGGCACTGCTCG 244  
Db 197 GCGGCTCGAGTACAGCGCGGAGCGGCTTCTGTCGCGCGCGGAGTGGCTGG 256

QY 245 CAGAATTCAACTCGCGCGCAACTGAAATCAAAAGCTCATGAACTCCAGAAAGGACAA 304  
Db 257 CGCTGCGGACCGCGCGCTTGCAGATGTCGGCGCGGACCGCGTGTGTCGCGCGCGC 316

QY 305 TCCGACTGCACTCATGCACTTCTTGGGCACTTGGATGGTCCCCCACTTATCCGACAT 364  
Db 317 TCGCGCTCGGCTGATGTCGACCGGCTGCGGTGACGTACCCGAGTTGTCGACGCGT 376

QY 365 TCCGCGCGAAACACCCCAACGTAGAAATTCRAACTCCACCAAGCGGCAAGCAATGCTCTCG 424  
Db 377 ACCGCGCGACGACCCGAGACGTCATGTCCTGCTCCGTTCCGGGCGGAGCGAGTGG 436

QY 425 TAGATCGTGTGTTGCTGATGAACTGACCTCGCAATTAGTTGGCCCCCAAACTGCGAGG 484  
Db 437 TGGCGCGATCCGGAACCGGAACTGGACATCGCCTTCTCGGTTGCGGAGGCGAAC 496

QY 485 TTGTGACTCTTTAGGTTGGGCGCACCTGCTTCGTCAGGACTTGCCTTAGCTGTTCGCG 544  
Db 497 GCGCACCCGCTGTGAAACCGTTGTTCTCGATCACGACGACGACTACTGTTGGTGGCGG 556

QY 545 CAGATCACCGCTTCCCTCTCTTCTGCGCAAGGAGAAATTGCGCTTGTATTACTGCGCGG 604  
Db 557 CCGGCGACCGGCTCGCGGTGTCTC-----CCAGGTACGCTGCGGAGATCGCGG 607

QY 605 AAGAACTTTCTGTCGCGATGCGAGCAGGTTTCGGACCCCGACTCTCTCATGATGATTAAG 664  
Db 608 AGGAGACGTACGTGGACTTCGTGGCGGAGACCCCGCGGCGCCAGTCCGATCAGGCGT 667

QY 665 CCGAAGAACCGGTTTGTTCCTCAATGTGTTTTCGAATCCATGGAATCACCACCGTTCG 724  
Db 668 TCGCGCTGCGGCGCTGTGTCGCGACGTCGCGCTACGAGGCGGCTGTGTCGAGCTGATCA 727

QY 725 CAGGCTTGTGACGCGCAGGTCCTCGGCGTGTGTTGTTCCGATGG 769  
Db 728 CCGGTTGATCACGCGCGGCTCGCGATCGCGCTGTCGCGCTGG 772

—

QY 401 ACCAAGGGGAGCAATGCTCTCTGGTAGATCGTGTGTTTGGCTGATGAACCTGACCTCGCAT 460  
Db 377 AGAGGCGCAACCGCGCGCTCGGTGAGATGCTGCGGAGGCGGACTCGCATGTGCGCG 436  
QY 461 TAGTTGGCCCAACCTGCGGAGTTGGTACCTCTTTTAGGGTGGGCGCCACTGCTTCGNC 520  
Db 437 TCGCTTCGGTACGAGGTGCGCGCGCGGCGGAGAAATGGAGCACTTGGTGTACGCC 496  
QY 521 AACGACTTGCCTAGCTGTTCGCGAGATCACCGGCTTGCCTCTTTTCTGGCCAAAGGAG 580  
Db 497 CCTGTCTACCGACCTGCTGTCGGCTGCTGTCGCCGAGGGCACCGCTGCGCGACCG 556  
QY 581 AATTGCGGTGATTACTGCGCGGAAGACCTTTCTGTCGCGATCGAGCAGGTTTCGGCA 640  
Db 557 A---GTCCGTCGCCATTCGCGGACCTTCGCCGCGACCCGTTGGATCGCGGATGCCCGCT 613  
QY 641 CCGGACTCTCATGATGATCATAGCCGAAGAGCGGTTTGTTCCTCAATGTGTTTCG 700  
Db 614 GCGCGGTGAGTCTGTCGAGGTGTGCGAGAGCGCGGCTTCACTCCCGGATCGACTTCG 673  
QY 701 AATCCATGGAATCACACCGTGCAGGCTTGTTCAGCGCAGGCTCTCGCGTGTGTTGG 760  
Db 674 CGACCGACGACTACCGCGGCTGTCGCTGTGTGAGCGCGGCTCTCGCGTGGCGGTAC 733  
QY 761 TTCCG 765  
Db 734 TGCCG 738

## RESULT 13

US-10-156-761-3218

; Sequence 3218, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 3218

; LENGTH: 1008

; TYPE: DNA

; ORGANISM: Streptomyces avermitilis

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1008)

US-10-156-761-3218

## Query Match

Best Local Similarity 6.5%; Score 60.4; DB 15; Length 1008;

Matches 190; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 1 GTGACATGGGCAACGACGCGGAGACCTCGCAATCGACGACTACGAGCTTCATTCA 60  
Db 7 GGGATCAGGACGAAAGGGGCGGCGCACCGACCTCGCCAGCTCGGGCTTCGCGGCC 66  
QY 61 GTGCTCAATCAGGGCCACTCACCGAAATGCGGAAAGATTAGGATCCCGCAGGCCACA 120  
Db 67 GTGCGGAGCATCTGACTTCGCGACGCGCGCGCGCGATCGGATGAGCCAGCCGCG 126  
QY 121 CTTTCCAGAGCAATCAGCGAGTGGAAACACGAGGACCCCACTTTTCCAGCGGCC 180

Db 127 CTGTGCGGCGCGGTCTCGCCCTTGAGAGAGACACTCGGGGTGACCCCTCTCGAGCGTACG 186  
QY 181 GGCCGCAAACTCGTCTCAACCAACGAGGCGCACGCTTCTCAACCAACGAGCGGCATC 240  
Db 187 ACCGCGNAGGTCTGCTCTCGCCCGCGGTGCGCGTCTCGCGTACGCGCCAGGCGGTG 246  
QY 241 GTGCGAGAAATCAACTCGCGCCGCAACTGAAATCAACGCTCATGGACCCAGAAAAGGC 300  
Db 247 CTCGACGAGTCCGGCGCGCTGATGAGGAGGCGGAGGCGGTACGCGCCCGTTTCCCGGG 306  
QY 301 ACAATCCGACTGACTTCAATGATTCCTTGGGCACTTGGATGGTCCCGCAACTTATCCGA 360  
Db 307 GCGCTGCGGCTCGGGGTGATCCCGACCGTCCGCGCGTATCTGCTGCCGACCGTCTCAGG 366  
QY 361 ACATTCGCGCGCAACACCCCAACGTAGAAATTCCAACTCCACCAAG 406  
Db 367 CTCGTCCACGACCGCTACCCCGACCTCGACCTCCAGGTCCACGAGG 412

## RESULT 14

US-10-156-761-1152

; Sequence 1152, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 1152

; LENGTH: 876

; TYPE: DNA

; ORGANISM: Streptomyces avermitilis

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(876)

US-10-156-761-1152

## Query Match

Best Local Similarity 6.5%; Score 60.2; DB 15; Length 876;

Matches 203; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

QY 29 TCGGAATCGAGACCTACGCGACTTCATTTCAGTCGCTCAATTCAGGCGACCTCACGAAA 88  
Db 2 TGGAACTGCGCACCTTCAGCATTTTCGTCGCGCTCGCGGAGGACGAGCACTTCACCGGG 61  
QY 89 CTGCGGAAGATTAGGCATCCGCGAGCCACACTTTCAGAGCAATCAGCGAGTGGAAA 148  
Db 62 CAGCGAAACGGCTGATGTGTGCGAGTCGGGCTGTCGCGCTCGATCCGCGCCCTGGAGC 121  
QY 149 AACCGCAGGCAACCCCACTTTTCGACCGCGCGCGCAAACTCGCTCTCAACCAACGAG 208  
Db 122 GGGAGCTCCAGGCGCGCTGTTGTCGCGCACCCACCGCAGTGTACGCTCACCGAGCCG 181  
QY 209 GCCAGCCTTCTCAACCAACGCGAGCGCATTCGTCGAGAAATTCAACTCCGCGCAACTG 268  
Db 182 GCGCGCGCTGTCGGGCGAGCGGAGCGGATTCCTGGCCAGGTCCGGGCGGCTCACGAGC 241  
QY 269 AAATCAACGCTCATGGAACCCAGAACCAAGGCAACATCCGACTGCACTTCATGCACTCT 328  
Db 242 CGGTGGCGCGCGTGCAGGCGCGTGTGTCGCGCGCACGCTCTCGCTGGGCGCCAGGAGTGCA 301

QY 329 TGGGCACTTGTGTGCTCCCGAATTATCCGAACATTCGCGCCGAAACACCCCAACGTAG 388  
Db |||||  
QY 302 TCGCGGGGTGATGTGGCCCGGCTGCTGCGCGGGTTCGCGGGGAGATCCGGATGTGG 361  
Db |||||  
QY 389 AATTCAACTCCACCAAGCGGAGCAATGCTCCTGGTAGATCGTGTGTTGGCTGATGAAA 448  
Db |||||  
QY 362 AGATCCGGTGTGCGCCAGGCGGGCTCGGGTGCCTGCGGAGGAAGTCCGCGCGGGCGTC 421  
Db |||||  
QY 449 CTGACCTCGCATTTAGTTGGCC 469  
Db |||||  
QY 422 TCGATCTGCGCCTTCGCGGTCC 442  
Db |||||

## RESULT 15

US-10-282-122A-31440  
; Sequence 31440, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Lianguo  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31440  
; LENGTH: 891  
; TYPE: DNA  
; ORGANISM: Pseudomonas putida  
US-10-282-122A-31440

Query Match 6.5%; Score 60; DB 17; Length 891;  
Best Local Similarity 46.4%; Pred. No. 1.2e-08;  
Matches 195; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

QY 38 ACGACCTACGAGCTTCACTTTTCAGTCGCTCAATCAGGCCACCTCACCGAAACTGCCGAAA 97  
Db |||||  
QY 17 AAGACATCAGATCTTTCGTGAAATCGCGAATCGGGAGTTTCATGAGCGCCCGGAAG 76  
Db |||||  
QY 98 GATTAGGCATCCCGCAGGCCACACTTTCAGACGAATCAGCGGAGTGGAAAAACACGCAG 157  
Db |||||

Db 77 CGTCCACTTGTCCAGCGGCCCTGAGCCGACGGATCAAGAAGCTGGAGGAAGSCCTGG 136  
QY 158 GCACCCCACTTTTTCGACCGCGCCGCGCAAACTCGTCTCTCAACCAACGAGGCCACGCCT 217  
Db |||||  
QY 137 GCACCTCGCTACTTGAACGCACTCGCCGGGTGAGCTTGCAGCGCTGAGCGCGCGGACT 196  
Db |||||  
QY 218 TCCTCAACACGCGCAGCGCCATCGTCGAGAAATTCAACTCCGCGCGCAACTGAAATCAAAC 277  
Db |||||  
QY 197 TCCTGCCCAAGCCAGCGGCTGCTGGATGATTTTGAAGACTCGATCTCAGCATCCCGG 256  
Db |||||  
QY 278 GCCTCATGGACCCAGAAAAAGSCAATCCGACTGGACTTCATGCAATTCCTTTGGSCACTT 337  
Db |||||  
QY 257 AGTGGCGAGCGCCAGACCGGTACCGCTCGCTGATTCCTACCGCGGGTCT 316  
Db |||||  
QY 338 GGATGGTCCCGCAACTTATCCGAACATTCGCGCGGGAACACCCCAACGTAGAAATCCAAAC 397  
Db |||||  
QY 317 ACTTCTCGCTGCTGATTCGCGGATTAACAACGAGCAATACCCGAAATTCGCGCATTCGCC 376  
Db |||||  
QY 398 TCCACCAAGCGCGCAGCAATGCTCCTGGTAGATCGTGTGCTGATGAAACTGACCTCG 457  
Db |||||  
QY 377 TGCTGGACCTTAGCGCCCAACGACGCGGCTCGAAGCGGTGCTGCGCGGCGAGGCCGACTTCG 436  
Db |||||

Search completed: April 13, 2005, 02:13:43  
Job time : 633 secs